AFOSR TR 96 0010

NorthWest Research Associates, Inc.
P.O. Box 3027 • Bellevue, WA 98009-3027

NWRA-CR-95-R147 13 November 1995

Gravity Waves in the Atmosphere: Instability, Saturation, and Transport

Final Technical Report
Covering 20 May 1992-19 May 1995

Prepared by
Timothy J. Dunkerton

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Prepared for
Maj. James T. Kroll
Air Force Office of Scientific Research
Bolling AFB, DC 20332-001

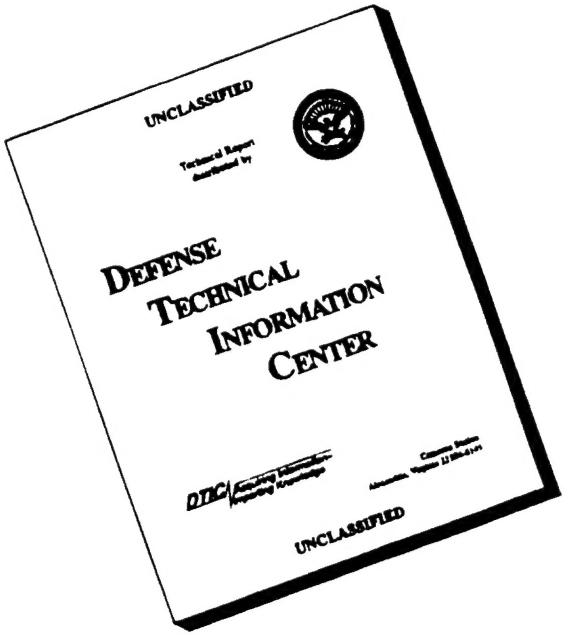
Contract No. F49620-92-C-0033

19960207 025

DTIC QUALITY INSPECTED 1

300-120th Ave. N.E. • Suite 220, Building 7 • Bellevue, WA 98005 • (206) 453-8141 • FAX (206) 646-9123

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or another aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	13 November 1995	Final Technical, 5/20/92 - 5/19/95

4. TITLE AND SUBTITLE

Gravity Waves in the Atmosphere: Instability, Saturation, and Transport

5. FUNDING NUMBERS

F49620-92-C-0033

6. AUTHOR(S)

Timothy J. Dunkerton

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Northwest Research Associates, Inc.
300 120th Ave NE, Bldg 7, Ste 220
P.O. Box 3027
Bellevue, WA 98009-3027

8. PERFORMING ORGANIZATION REPORT NUMBER

NWRA-CR-95-R147

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Office of Scientific Research
110 Duncan Avenue, Suite B115
Bolling AFB, DC 20332-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The excitation, propagation, breakdown, absorption of gravity waves in the atmosphere is fundamentally important to the momentum balance, transport of heat and constituents, and general circulation. This report describes the results of several studies pertaining to excitation of inertia-gravity waves in midlatitude baroclinic systems, effects of the quasi-biennial oscillation on the inertia-gravity and Kelvin wave spectrum at the equator, the role of gravity-wave drag in the extratropical QBO, destabilization of large-scale tropical waves by deep moist convection, and a general theory of equatorial inertial instability on a zonally nonuniform, nonparallel flow.

14. SUBJECT TERMS

inertia-gravity waves, inertial instability

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
U	U	U	

Table of Contents

Inertial Instability of Nonparallel Flow on an Equatorial β Plane 1

Seasonal Development of the Extratropical QBO in a Numerical Model of the Middle Atmosphere 17

CISK and Evaporation-Wind Feedback with Conditional Heating on an Equatorial Beta-Plane 35

Horizontal Buoyancy Flux of Internal Gravity Waves in Vertical Shear 45

Generation of Inertia-Gravity Waves in a Simulated Lifecycle of Baroclinic Instability 57

Reprinted from JOURNAL OF THE ATMOSPHERIC SCIENCES, Vol. 50, No. 16, 15 August 1993
American Meteorological Society

Inertial Instability of Nonparallel Flow on an Equatorial β Plane

TIMOTHY J. DUNKERTON

Inertial Instability of Nonparallel Flow on an Equatorial β Plane

TIMOTHY J. DUNKERTON

Northwest Research Associates, Bellevue, Washington

(Manuscript received 22 June 1992, in final form 7 December 1992)

ABSTRACT

A simple theoretical model was developed to investigate the inertial instability of zonally nonuniform, nonparallel flow near the equator. The basic state was independent of height and time but included cross-equatorial shear with longitudinal variation, as observed in the tropical mesosphere and elsewhere. Numerical solutions were obtained for the most unstable modes.

It is shown that, in addition to previously known "global" (symmetric and nonsymmetric) modes of inertial instability, there exist "local" modes within regions of anomalous potential vorticity. Local modes may be exactly stationary or display zonal phase propagation, but are distinguished from global modes by their zero group velocity and concentration of amplitude within, or downstream from, the region of most unstable flow. Local stationary instability has the largest growth rate and occurs in strong inhomogeneous shear when the in situ mean flow is near zero, that is, quasi-stationary with respect to the (stationary) basic-state pattern. This situation is expected in an equatorial Rossby wave critical layer.

The local mode has properties similar to those of "absolute" instability of nonparallel flow as discussed elsewhere in fluid dynamics.

1. Introduction

Inertial instability may arise in conservative axisymmetric flow near the equator when there is nonzero latitudinal shear $\gamma = \bar{u}_y$. In this case potential vorticity (PV) on one side of the equator is anomalous, satisfying a necessary condition for centrifugal parcel instability (Dunkerton 1981). Analytic eigenmodes of symmetric instability in zonally uniform shear on an equatorial β plane were obtained by Dunkerton and independently by Stevens (1983). Subsequently Boyd and Christidis (1982) and Dunkerton (1983) found that low-wavenumber, zonally nonsymmetric instabilities have larger growth rates than symmetric instability when (i) the vertical wavenumber m is near or below a neutral point of symmetric instability

$$\frac{|m|}{N} = \epsilon_{\text{neut}}^{1/2} = \frac{4\beta}{\gamma^2} \tag{1.1}$$

(N is static stability and $\beta = 2\Omega/a = 2.29 \times 10^{-11}\ \text{m}^{-1}\ \text{s}^{-1}$), and (ii) nonzero *integer* zonal wavenumbers $s = ka$ exist below a short-wave cutoff $k_{\max} \equiv \beta/\gamma$. Although symmetric instability has maximum inviscid growth approaching $\gamma/2$ as $|m| \to \infty$, the nonsymmetric mode is preferred when second-order viscosity and diffusion are added and conditions (i) and (ii) are met (Dunkerton 1983). Similar conclusions pertain to midlatitude flow containing nonsymmetric inertial and barotropic instabilities, as shown by Stevens and Ciesielski (1986).

Anomalous PV exists in the tropical winter mesosphere (Dunkerton 1981; Hitchman and Leovy 1986), upper troposphere between South America and the Indian Ocean (Liebmann 1987), lower troposphere near Indonesia (Krishnamurti et al. 1985, 1988), and Pacific south equatorial current (Philander 1989, p. 62). Inertial instabilities were found in a nearly inviscid 2D axisymmetric model of the troposphere (Held and Hou 1980) and in 3D middle atmosphere GCMs (Hunt 1981; O'Sullivan and Hitchman 1992). This instability was suggested as a possible explanation of layered structures in the tropical mesosphere (Hitchman et al. 1987; Fritts et al. 1992) and mesoscale anomalies in the midlatitude upper troposphere (Ciesielski et al. 1989). Conditional symmetric instability is thought to explain frontal rainbands according to Bennetts and Hoskins (1979) and several others.

Unfortunately for the theory, none of the regions of anomalous PV in atmosphere or ocean is zonally symmetric. Observations and numerical models suggest, for example, that inertial instabilities in the mesosphere coincide with tropical penetration of planetary Rossby waves from midlatitudes (Hitchman et al. 1987; O'Sullivan and Hitchman 1992). For the theory to be relevant, it must be generalized to a nonparallel or zonally nonuniform basic state.

The purpose of this paper is to develop a simple theoretical formalism and thereby to document two types of inertial instability in nonparallel flow: local instability (stationary and zonally propagating) and

Corresponding author address: Dr. Timothy J. Dunkerton, Northwest Research Associates, P.O. Box 3027, Bellevue, WA 98009.

© 1993 American Meteorological Society

global instability (symmetric and nonsymmetric). Of these, the local instabilities are new and probably most relevant to middle atmosphere models in which breaking planetary waves produce regions of significantly anomalous PV near the equator.

The theoretical model is described in section 2, including a brief review of symmetric and nonsymmetric instability. Growth rates and structure of nonparallel instabilities are illustrated in section 3 for a wide range of model parameters using a simple analytic basic state resembling the onset of planetary wave breaking. The theoretical interpretation of local instability is discussed further in section 4 and the Appendix.

2. Background

To isolate the effect of horizontal basic-state variations, a simple two-dimensional (x–y) model is desirable, analogous to the shallow-water system. In a vertically continuous atmosphere, these equations describe perturbations oscillating in height with constant sinusoidal variation. Vertical wavelength is then contained in Lamb's parameter ϵ (Andrews et al. 1987) and is specified a priori. In this section, the simple model is developed, its solution method described, and a brief review given of parallel flow instabilities (several features of which carry over to the nonparallel case).

a. Theoretical model

The hydrostatic primitive equations on an equatorial β plane are

$$u_t + uu_x + v(u_y - \beta y) + wu_z + \phi_x = X \quad (2.1a)$$

$$v_t + u(v_x + \beta y) + vv_y + wv_z + \phi_y = Y \quad (2.1b)$$

$$\phi_{zt} + u\phi_{zx} + v\phi_{zy} + w(N^2 + \phi_{zz}) = Q \quad (2.1c)$$

$$u_x + v_y + \frac{1}{\rho_0}(\rho_0 w)_z = 0 \quad (2.1d)$$

where u, v, w are zonal, meridional, and vertical velocity, ϕ is geopotential, ρ_0 is basic-state density $\rho_s \exp(-z/H)$, and X, Y, Q are dissipative terms. Equations (2.1) are essentially those of Andrews and McIntyre (1976) written in log-pressure notation of Holton (1975).

To formulate an idealized model of instability the dependent variables may be expanded as

$$u(x, y, z, t) = \tilde{U}(x, y) + u'(x, y, z, t) \quad (2.2)$$

and similarly for v, w, ϕ. Here $\tilde{U}(x, y)$ is a barotropic basic state independent of height and time for the purpose of linear stability analysis. Linearized perturbation equations are then

$$u'_t + \bar{u}u'_x + v'(\bar{u}_y - \beta y) + \phi'_x = A + K\nabla_H^2 u' - \nu m^2 u' \quad (2.3a)$$

$$v'_t + \bar{u}v'_x + \beta y u' + \phi'_y = B + K\nabla_H^2 v' - \nu m^2 v' \quad (2.3b)$$

$$\phi'_t + \bar{u}\phi'_x + \epsilon^{-1}(u'_x + v'_y) = C + K\nabla_H^2 \phi' - \nu m^2 \phi' \quad (2.3c)$$

under the following assumptions:

1) The basic-state velocity field is barotropic, temporally constant, and horizontal ($\tilde{W} \equiv 0$). More generally we could regard the mean flow as slowly varying in height and time (Boyd 1978).

2) Advection of basic-state temperature is neglected, that is, $\mathbf{u}' \cdot \nabla_H \Phi_z = 0$. This term has a counterpart $\mathbf{u}' \cdot \nabla_H \Phi$ normally retained in the shallow-water system.[1]

3) Perturbations vary as $\exp imz$, and $m^2 \gg 1/(4H^2)$ whereupon $\epsilon \approx m^2/N^2$ as in (1.1).

4) $|\phi_{zz}| \ll N^2$.

5) Dissipative terms are written as second-order viscosity and diffusion.

In (2.3a–c) the zonally symmetric component of $\tilde{U}$ was kept on the lhs, that is,

$$\tilde{U} \to \bar{u}(y) + U(x, y) \quad (2.4a)$$

$$\tilde{V} \to V(x, y) \quad (2.4b)$$

and on the rhs

$$-A \equiv (Uu'_x + Vu'_y) + (u'U_x + v'U_y) \quad (2.5a)$$

$$-B \equiv (Uv'_x + Vv'_y) + (u'V_x + v'V_y) \quad (2.5b)$$

$$-C \equiv (U\phi'_x + V\phi'_y). \quad (2.5c)$$

b. Solution method

To solve (2.3a–c) a semi-implicit method was used in which variables were expanded in zonal harmonics (letting k be integers for the moment),

$$\begin{Bmatrix} u' \\ v' \\ \phi' \end{Bmatrix} \equiv \frac{1}{2} \sum_{k=-\infty}^{\infty} \begin{Bmatrix} u_k \\ iv_k \\ \phi_k \end{Bmatrix} \exp \frac{ikx}{a} \quad (2.6a)$$

$$\begin{Bmatrix} U \\ V \end{Bmatrix} \equiv \frac{1}{2} \sum_{k \neq 0} \begin{Bmatrix} U_k \\ V_k \end{Bmatrix} \exp \frac{ikx}{a} \quad (2.6b)$$

$$\begin{Bmatrix} A \\ B \\ C \end{Bmatrix} \equiv \frac{1}{2} \sum_{k=-\infty}^{\infty} \begin{Bmatrix} A_k \\ B_k \\ C_k \end{Bmatrix} \exp \frac{ikx}{a}. \quad (2.6c)$$

Terms on the lhs of (2.3a–c) were treated implicitly in time, that is,

[1] In real atmospheres the basic state will include variations of static stability induced by potential vorticity anomalies (Hoskins et al. 1985). For simplicity, only the variation of induced horizontal velocity U, V was retained in the definition of A, B, C. This assumption is reasonable in the tropical middle atmosphere, where incident Rossby waves have approximately horizontal ray paths.

$$\hat{u}_k = \tfrac{1}{2}(u_k^n + u_k^{n+1}), \quad \text{etc.,} \tag{2.7}$$

so that for each harmonic

$$-\hat{\omega}\hat{u}_k + \hat{v}_k(\bar{u}_y - \beta y) + \frac{k}{a}\hat{\phi}_k = -i\tilde{A}_k^n - \omega u_k^n \tag{2.8a}$$

$$+\hat{\omega}\hat{v}_k + \beta y\hat{u}_k + \frac{\partial}{\partial y}\hat{\phi}_k = \tilde{B}_k^n + \omega v_k^n \tag{2.8b}$$

$$-\hat{\omega}\hat{\phi}_k + \epsilon^{-1}\left(\frac{k}{a}\hat{u}_k + \frac{\partial}{\partial y}\hat{v}_k\right) = -i\tilde{C}_k^n - \omega\phi_k^n, \tag{2.8c}$$

where $\hat{\omega} \equiv \omega - k\bar{u}/a$, $\omega \equiv 2i/\Delta t$, and

$$\tilde{A}_k = A_k + \text{diffusion terms, etc.} \tag{2.9}$$

After some manipulation (2.8a–c) reduce to a geopotential equation

$$\Delta\frac{\partial}{\partial y}\left(\frac{\phi_y}{\Delta}\right) - \phi\left[k^2 + \frac{k\beta}{\hat{\omega}\Delta}(y\Delta_y - \Delta)\right] = \epsilon\Delta\phi + \mathcal{D} \tag{2.10}$$

for each harmonic, where

$$\Delta \equiv \beta y(\beta y - \gamma) - \hat{\omega}^2. \tag{2.11}$$

In (2.10) $\mathcal{D}$, whose definition is omitted for brevity, includes the explicit forcing and solutions from the previous time step. Equation (2.10) was solved by discretizing in y with 64 grid points and using a tridiagonal algorithm. The domain was periodic in x and extended from $-4y_s$ to $+6y_s$ where $y_s \equiv \gamma/2\beta$. There are no critical latitude ($\hat{\omega} = 0$) or inertial latitude ($\Delta = 0$) singularities when $\hat{\omega}$ is complex. A spectral rather than pseudospectral method was used to evaluate $\mathcal{D}$ since in this case U_k, $V_k \equiv 0$ for any $|k| \neq 1$; that is, a purely wave 1 contribution to basic-state variation in x was assumed. Product terms A_k, etc., could be determined efficiently without transforms. In the numerical integrations

$$\bar{u} = \gamma(y - y_c) + \bar{u}_0 \tag{2.12}$$

where γ and $\bar{u}_0$ are constants. Fields of U, V are specified in section 3.

c. *Instability of uniform parallel flow*

When $U = V = 0$, harmonic perturbations proportional to $\exp i(kx - \omega t)$ satisfy

$$-\hat{\omega}u + v(\gamma - \beta y) + k\phi = 0 \tag{2.13a}$$

$$+\hat{\omega}v + \beta y u + \phi_y = 0 \tag{2.13b}$$

$$-\hat{\omega}\epsilon\phi + ku + v_y = 0 \tag{2.13c}$$

where $\hat{\omega} = \omega - k\bar{u}$. The geopotential equation derived from (2.13a–c) is just (2.10) with $\mathcal{D} \equiv 0$ and $\hat{\omega}$ redefined as intrinsic frequency. This equation is linear in the "eigenvalue" ϵ. Boyd (1978) derived the meridional velocity equation

$$v_{yy} + \frac{2\epsilon\hat{\omega}k\gamma}{\epsilon\hat{\omega}^2 - k^2}v_y - v\left[\frac{k}{\hat{\omega}}(\beta - \gamma_y) + k^2 + \epsilon\Delta + \frac{2\epsilon k^2\gamma(\beta y - \gamma)}{\epsilon\hat{\omega}^2 - k^2}\right] = 0 \tag{2.14}$$

which is nonlinear in ϵ unless, for example, $\epsilon\hat{\omega}^2 \gg k^2$ in which case

$$v_{yy} + \frac{2k\gamma}{\hat{\omega}}v_y - v\left[\frac{k}{\hat{\omega}}(\beta - \gamma_y) + \epsilon\Delta\right] \approx 0. \tag{2.15}$$

If $k = 0$

$$v_{yy} - \epsilon v[\beta y(\beta y - \gamma) - \omega^2] = 0 \tag{2.16}$$

and the eigencondition for symmetric instability is

$$\omega^2 + \gamma^2/4 = \frac{(2n+1)N\beta}{|m|} \tag{2.17}$$

(Dunkerton 1981). Solutions for v are Hermite functions (polynomials times a Gaussian) centered about a "shifted equator" $y_s \equiv \gamma/2\beta$, the center of anomalous vorticity region. For $n = 0$ the neutral point of symmetric instability (hereafter, "neutral point") is given by (1.1). For small k and $\gamma_y = 0$,

$$\omega^2 + \gamma^2/4 \approx \frac{N\beta}{|m|} + \frac{k\beta}{\omega\epsilon}. \tag{2.18a}$$

At $\epsilon = \epsilon_{\text{neut}}$,

$$\omega^3 \approx \frac{k\beta}{\epsilon}. \tag{2.18b}$$

Equations (2.18a,b) were derived in nondimensional form by Boyd and Christidis (1982).

At finite k the eigenproblem (2.10) or (2.14) must be solved numerically. Growth rates for $\gamma = 10^{-5}\ \text{s}^{-1}$ are shown in Fig. 1 illustrating the relative importance of symmetric and nonsymmetric modes. Curves were obtained from the geopotential equation by a shooting method; symbols correspond to values from the time-dependent model. They are in excellent agreement except at very small growth rates, where the time-dependent model could not determine growth rate accurately, and at large ϵ, where latitudinal resolution was inadequate.

An approximate model derived from (2.15) contains eigenfrequencies similar to the exact values of Fig. 1 over a large portion of parameter space. This model begins with constant γ and assumes that y variations of intrinsic frequency can be neglected outside Δ:

$$v_{yy} + \frac{2k\gamma}{\omega_0}v_y - v\left[\frac{k\beta}{\omega_0} + \epsilon\Delta\right] \approx 0 \tag{2.19}$$

where ω_0 is a constant (complex) intrinsic frequency within the unstable region. It follows that

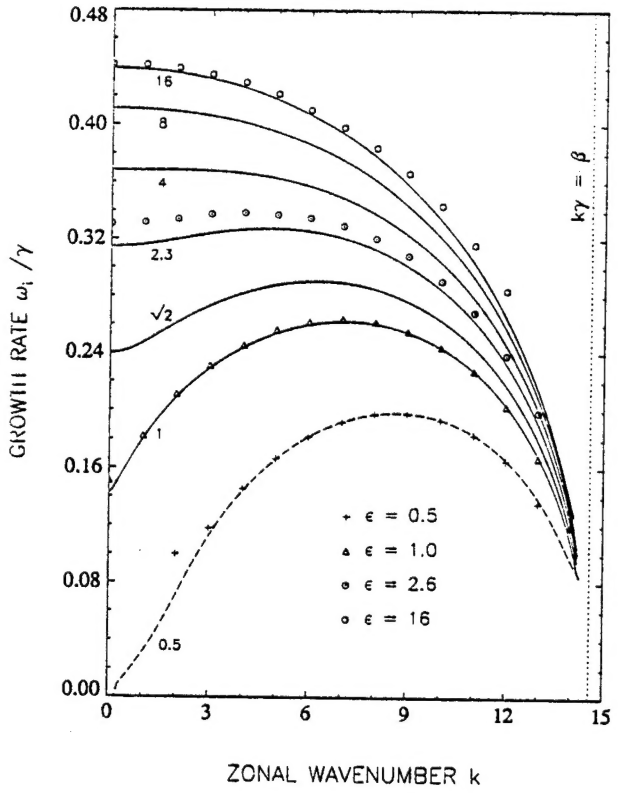


FIG. 1. Growth rate as a function of zonal wavenumber in parallel flow, for $\gamma = 10^{-5}\ s^{-1}$; ϵ in units $m^{-2}\ s^2$. Symbols denote results of time-dependent numerical model; solid and dashed curves show analytic results from a shooting method. (The dashed curve had no symmetric instability because $\epsilon < \epsilon_{neut}$.)

$$\frac{(\omega - k\gamma^2/\beta)^2}{1 - k^2\gamma^2/\beta^2} + \frac{1}{4}\gamma^2 = \frac{N\beta}{|m|}\sqrt{1 - k^2\gamma^2/\beta^2} + \frac{k\beta}{\omega_0\epsilon} \tag{2.20}$$

where

$$\omega_0 \equiv \omega - k\gamma^2\frac{\beta - 2\omega k}{2(\beta^2 - k^2\gamma^2)}. \tag{2.21}$$

A formal justification of approximations leading to (2.20) will not be given except to note that terms of $O(k^2\gamma^2)$ arise from Δ at large y and cannot be ignored for latitudinal trapping and quantization of eigenfunctions near the short-wave cutoff. The approximate dispersion relation (2.20) has the desired behavior at small and large k. In between, agreement with exact growth rates and phase speeds is good for $\epsilon \geq \epsilon_{neut}$ (not shown). At smaller ϵ the cubic equation fails in two ways: growth rates are overestimated at intermediate k and there is no Kelvin wave instability at small k below the neutral point (Boyd and Christidis 1982).

3. Instability of nonparallel flow

The time-dependent model of section 2 was used to determine the growth of unstable modes in nonparallel flow. Some characteristics of these modes are now described. The basic state, independent of height and time, was designed as a simple analog of incipient planetary wave breaking in the tropical mesosphere. The evolution of inertial instability in a time-dependent Rossby wave critical layer was recently discussed by O'Sullivan and Hitchman (1992) in a three-dimensional model. In order to understand the selection of unstable modes, it will be worthwhile to explore the parameter dependence of inertial instability in a simpler two-dimensional model.

a. Basic state

The basic state was defined by

$$\bar{u}(y) + U(x, y) = \gamma(y - y_c) - \frac{\gamma\psi y_s}{2}\left[1 + \tanh\left(\frac{y - y_c}{y_s}\right)\right]\cos(x/a) + \bar{u}_0. \tag{3.1}$$

In (3.1), the latitude of maximum shear is y_c, ψ is dimensionless amplitude of basic-state variation, $y_s = \gamma/2\beta$ is the center of anomalous zonal-mean vorticity as defined previously, and $\bar{u}_0$ is a constant mean flow independent of x, y. The profile (3.1) is shown in Fig. 2 for $y_c = 1.8y_s$, $\psi = 1$, and $\bar{u}_0 = 0$.

In Fig. 2, latitudinal shear at $y = y_c$ was increased (decreased) relative to the zonal mean by 50% at 180° (0°) longitude. The flow was inertially unstable at all x, but more so at the center of figure. This region could represent a zone of enhanced latitudinal shear set up by Rossby waves penetrating from latitudes north of the equator. In reality, there would be some latitudinal phase tilt during critical layer development (O'Sullivan and Hitchman 1992)—an unnecessary complication for stability analysis.

The meridional component of basic state $V(x, y)$ was included in all simulations, that is,

$$U_x + V_y = 0. \tag{3.2}$$

This had a minor effect compared to the zonal component (3.1). Although the details are omitted, it can be stated that all of the results shown here were similar when $V \equiv 0$. Local instabilities were also found when $U \equiv 0$ and $V = V(x)$ (localization due to $\partial V/\partial x$). The effect of V was generally to disrupt rather than enhance the instability in comparison to parallel flow. These cases were felt to be less interesting and were not investigated further.

b. Latitude of maximum shear

The latitude of maximum latitudinal shear was y_c according to (3.1). On the other hand, the maximum rate of "parcel instability" is $R \equiv \sqrt{-\beta y(\beta y + \zeta)}$ where ζ is relative vorticity. This quantity is contoured in Fig. 2; it did not always maximize at $y = y_c$ but was a function of y_c (and other parameters), having (for $\psi = 1$)

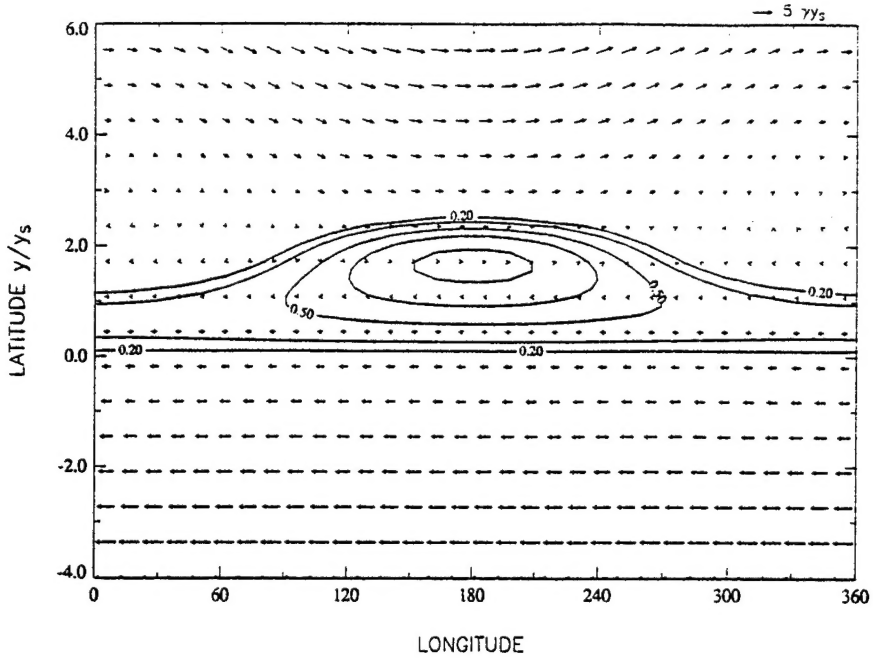


FIG. 2. Profile of nonparallel basic state (3.1) with $\psi = 1$. Contours show dimensionless parcel displacement growth rate R/γ. Contour interval 0.1, with 0.1 and 0.4 contours omitted.

an overall maximum value $R = 0.75\gamma$ when $y = y_c = 1.5y_s$. Maximum R was located north of y_s because the wave's contribution to basic state expanded the zone of instability at 180° (recall that y_s was the center of anomalous *zonal-mean* vorticity). As $|y_c| \to \infty$, however, the relative vorticity at $|y_c|$ became small relative to βy, so the region of anomalous vorticity next to the equator returned to a parallel configuration. For parallel flow the maximum $R = \gamma/2$ at $y = y_s$.

Growth rates of the most unstable mode as a function of y_c are shown in Fig. 3 for $\gamma = 4 \times 10^{-5}\ s^{-1}$, $\epsilon^{-1} = 256\ m^2\ s^{-2}$, $\psi = 1$, and $\bar{u}_0 = 0$. For these calculations eight zonal harmonics were used, and $\Delta t = 900$ s. Vertical wavelength ~5 km for this choice of ϵ when $N = 0.02\ s^{-1}$. The asymptotic forms of mode and growth rate were realized after a few days in most cases, although as will be seen later, this time scale depended on γ. The results appear reasonable: growth rate peaked at $y_c \approx 1.7y_s$ and for large $|y_c|$ instabilities developed as in parallel flow (not shown). The interpretation of Fig. 3 is slightly more complicated, however, because: 1) instabilities near the center of the plot were exactly stationary "s" and grew about twice as fast as in parallel flow—a larger variation than expected from R alone. Peak growth occurred a little to the right of $1.5y_s$. 2) Instabilities adjacent to those labeled "s" were zonally propagating but not global in extent.

The basic-state zonal flow $[\bar{u}(y) + U(x, y)$ in (3.1)] at location of maximum $R(180°, y)$ also varied with y_c; it was zero at this point when $y_c \approx 2.2y_s$. This variation affected the growth rate in addition to the variation of R. Growth was optimized for a combination that maximized R and minimized the in situ mean flow. The dependence of growth rate on mean flow speed, or breakdown of Galilean invariance, is characteristic of local instability of nonparallel flow as discussed, for example, by Pierrehumbert (1984). We return to this point in the next subsection and in section 4.

Instability structure for $y_c = 1.8y_s$ is shown in Fig. 4. The disturbance was exactly stationary and confined

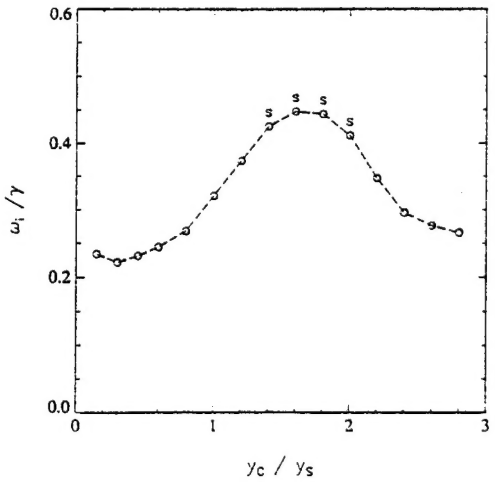


FIG. 3. Growth rate as a function of y_c. Local stationary instability denoted by "s."

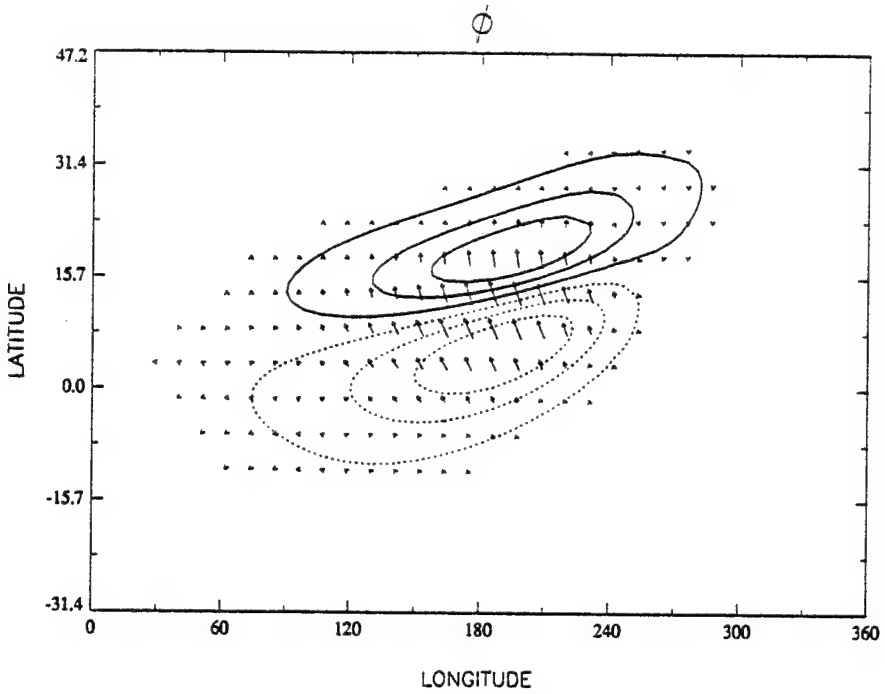


FIG. 4. Structure of local stationary mode for $\gamma = 4 \times 10^{-5}\ s^{-1}$, $\epsilon^{-1} = 256\ m^2\ s^{-2}$, $\psi = 1$, $\bar{u}_0 = 0$, and $y_c = 1.8 y_s$. Negative ϕ contours are dashed.

to the middle half of the domain. Velocity was directed from low to high pressure indicating a dynamical source of instability (Coriolis force). Curiously the geopotential phase tilted slightly although no tilt was imposed on the basic state (3.1). Identical results were obtained with doubled resolution in x, y and $\Delta t = 450$ s.

Figure 4 exemplifies "local stationary" instability. Note that 1) although localized in x it is not infinitesimal in size. This indicates that terms in addition to $v'U_y$ are important in the dynamics, as expected from section 2c. An oversimplified model could be formulated in which $v'U_y$ was the only term in A, all linear terms $\propto \partial/\partial x$ were dropped, and B, $C \equiv 0$. Growth rate in this case could be obtained from the symmetric formula (2.18a with $k = 0$) but was a function of x, implying a δ-function catastrophe. (This behavior was verified numerically.) Further, 2) absence of zonal propagation over a finite range of parameters (in this case, y_c) apparently requires the zonally symmetric ($s = 0$) component of perturbation, an important part of the total wave field. Most of the perturbation energy was contained in the lowest four harmonics. When the $s = 0$ component was artificially excluded from the numerical model, real phase speed was nonzero except where its trajectory crossed the $c_r = 0$ axis at a point near $y_c = 1.8 y_s$.

In parallel flow, nonsymmetric instability propagates zonally except in the special case $c_r(k, \epsilon) \equiv -\bar{u}_0$ (Dunkerton 1983; Stevens and Ciesielski 1986). Only symmetric instability is trivially "stationary" regardless of parameters. In nonparallel flow, local stationary instability is possible. It is in a sense locally symmetric, although this description ignores the spatial variation of instability structure (e.g., Fig. 4).

c. *Dependence on* $\bar{u}_0$

The term $\bar{u}_0$ in (3.1) added a constant mean flow independent of x, y. In parallel flow $\bar{u}_0$ would have no effect other than to shift phase speed by this constant.

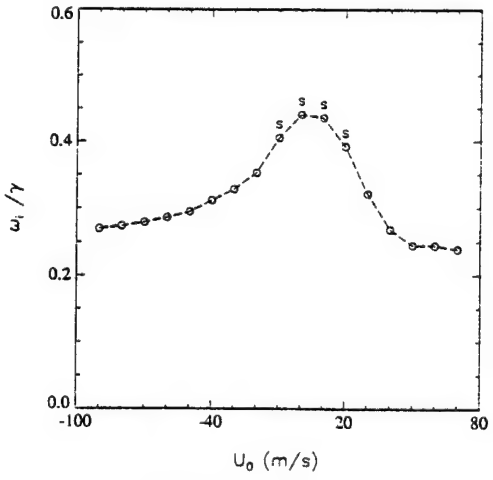


FIG. 5. Growth rate as a function of $\bar{u}_0$.

For local modes in nonparallel flow, growth rate is altered.

Figure 5 shows ω_i/γ for the same parameters as Fig. 2 but with $y_c = 1.8\,y_s$ and $\bar{u}_0$ varied over a large range. Local stationary instability occurred near the point of zero mean flow, but for large $|\bar{u}_0|$, the instability was displaced off center and forced to propagate in the direction of $\bar{u}_0$.

An example of this zonally propagating or advected instability is shown in Fig. 6 for $\bar{u}_0 = -30$ m s^{-1}. Propagation was to the west, and the pattern was observed to repeat every six days. Obviously this was insufficient time for global traverse, and in any case, the disturbance reentering from the right had little role to play in the sequence.

It is thought that local instabilities generally belong to a class defined by "absolute" instability (Pierrehumbert 1984). Absolute instability is sensitive to a mean flow that displaces the center of mode away from the source of instability and thereby lowers the effective growth rate. When the mean flow is stable at large $|x|$, a sufficiently large $|\bar{u}_0|$ can stabilize the problem. This did not happen at the edges of Fig. 5 because of the zonally symmetric component of shear and periodic boundary conditions; advected modes simply wrapped around and continued to exist as global modes in x, with smaller growth rate. A sponge region near $x = 0$, to simulate an infinite domain, dramatically altered the advected mode as discussed at the end of this section. The interpretation of local modes as absolute instability is deferred to the Appendix.

Of the two subtypes of local instability (stationary and advected), the stationary variety seems more relevant to equatorial Rossby wave critical layers, for instance, when a quasi-stationary Rossby wave induces a region of significantly anomalous potential vorticity

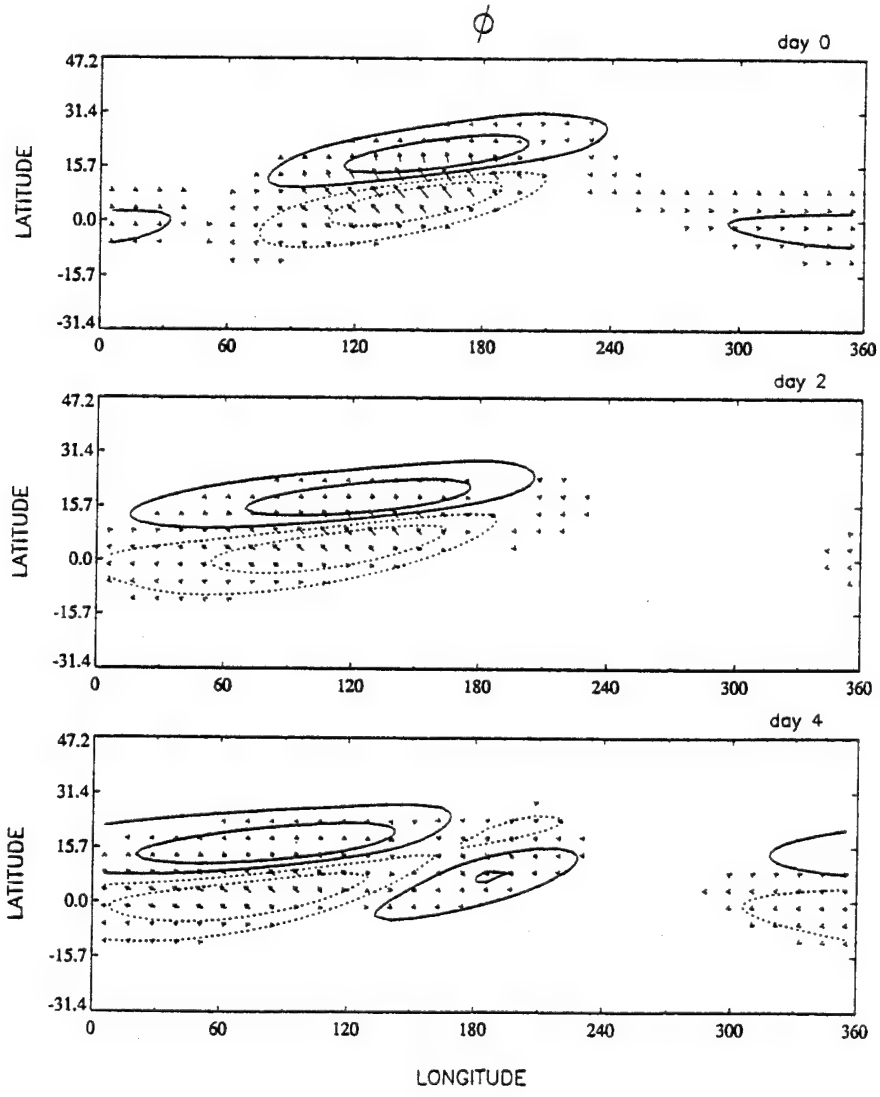


FIG. 6. Structure of unstable advected mode for parameters of Fig. 4 but with $\bar{u}_0 = -30$ m s^{-1}.

near the zero-wind line. Here the advecting mean flow is small. That is presumably why instabilities simulated by O'Sullivan and Hitchman (1992) were quasi-stationary (locally symmetric) and followed the anomalous PV during critical layer development.

d. *Transition to longitudinally global instability*

The value of ψ in (3.1) determined the zonal variation of shear. Not surprisingly, it was found that as ψ decreased there was a transition to global instability, as shown in Fig. 7 (p signifying a globally propagating or parallel-flow instability). Growth rate was reduced as $\psi \to 0$. Parameters were the same as before, but with $y_c = 1.8y_s$, $\bar{u}_0 = 0$, and variable ψ. The structure became essentially that of nonsymmetric instability in parallel flow (Stevens and Ciesielski 1986), with enhanced amplitude near the most unstable region (see Fig. 8, for $\psi = 0.2$). These were global modes that recycled through the entire domain. (Following Pierrehumbert, "global" refers only to the *longitudinal* extent of eigenmodes.)

Results at very small ψ depended on initial conditions. A wavenumber 1 initial disturbance, used in most simulations here, produced wave 1 nonsymmetric instability. Initial wave 2 led to wave 2 instability at small ψ, with slightly larger growth rate. Wave 3 was near the short-wave cutoff and less unstable.

The time-dependent model simulated only the most unstable mode; however, we expect local and global modes of instability to coexist, at least in cases such as Fig. 2 where the flow is unstable at all x. Whether global modes of inertial instability are important in the middle atmosphere or anywhere else is uncertain.[2]

e. *Variation of* ϵ

The value of ϵ^{-1} in Figs. 3–8 (256 $m^2 s^{-2}$) was near the neutral point $\epsilon^{-1}_{neut} = 305\ m^2 s^{-2}$ for $\gamma = 4 \times 10^{-5}\ s^{-1}$. Vertical wavelength was ~5 km when $N = 0.02\ s^{-1}$. [The reader may refer to (1.1) for other N, and recall the transformation of variables in Dunkerton (1983) allowing generalization to arbitrary values of shear.]

Recalling the discussion of section 2c, the neutral point limits symmetric instability in parallel flow. There is no *neutral curve* at small ϵ (except at small k adjacent to the neutral point; see Boyd and Christidis 1982) to prevent instability at nonzero $s = ka$. For $\gamma = 4 \times 10^{-5}\ s^{-1}$, several unstable integer wavenumbers are allowed below the short-wave cutoff ($s = 0$–3). Growth rate varies monotonically, increasing with ϵ to a limiting value $\gamma/2$ (cf. Fig. 1).

In nonparallel flow, the growth of local instability depended on ϵ in a similar way as shown in Fig. 9. The same parameters were used in (3.1) with $y_c = 1.8y_s$, $\psi = 1$, and $\bar{u}_0 = 0$. Growth rate varied monotonically, reaching a plateau at higher ϵ. Over most of this range the instability was stationary, displaying slow eastward propagation only at small ϵ. It seems clear from this example that nonparallel instability exists below the neutral point whether one adopts a zonal-mean flow value ($\epsilon^{-1}_{neut} \sim 305\ m^2 s^{-2}$) or hypothetical local value ($\epsilon^{-1}_{neut} \sim 1545\ m^2 s^{-2}$) at center of domain.

The structure of instability at small ϵ had several interesting features (not shown): the latitudinal scale was enlarged, as in classical equatorial wave theory, so instability extended into the stable part of the domain. Its structure resembled a Rossby wave in the Northern Hemisphere and eastward inertia–gravity wave in the Southern Hemisphere (Matsuno 1966). The real phase speed, though slow eastward, was for practical purposes stationary with respect to the ground. The disturbance propagated, with respect to the fluid, westward north of the equator and eastward south of the equator. Because of latitudinal shear, a fused Rossby/inertia–gravity structure could be excited by equatorial instability that acted as a kind of "wavemaker" in this instance. [Frontal excitation of inertia–gravity waves in a localized region of conditional symmetric instability was noted by Thorpe and Rotunno (1989) and Jones and Thorpe (1992).] The instability was, however, local in x.

The structure of instability at large ϵ resembled Fig. 4 but was contracted in latitude, remaining local in x.

Figure 9 suggests that as in parallel flow, scale selection at finite $|m|$ is unrealized in the simple model.

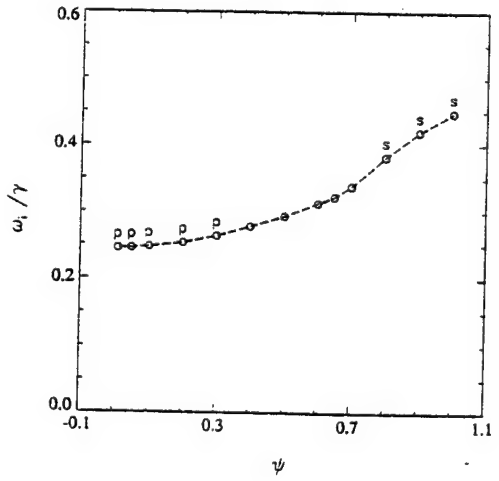


FIG. 7. Growth rate as a function of ψ. Globally propagating or parallel-flow instability denoted by "p."

[2] There is an intriguing visible light photograph of Jupiter taken by the Hubble Space Telescope on 28 May 1991 revealing horizontally tilted, banded structures next to the equator (e.g., see the cover page of AGU publication *Earth in Space*, January 1992 issue). This would suggest horizontally divergent circulations, perhaps due to a dynamical instability.

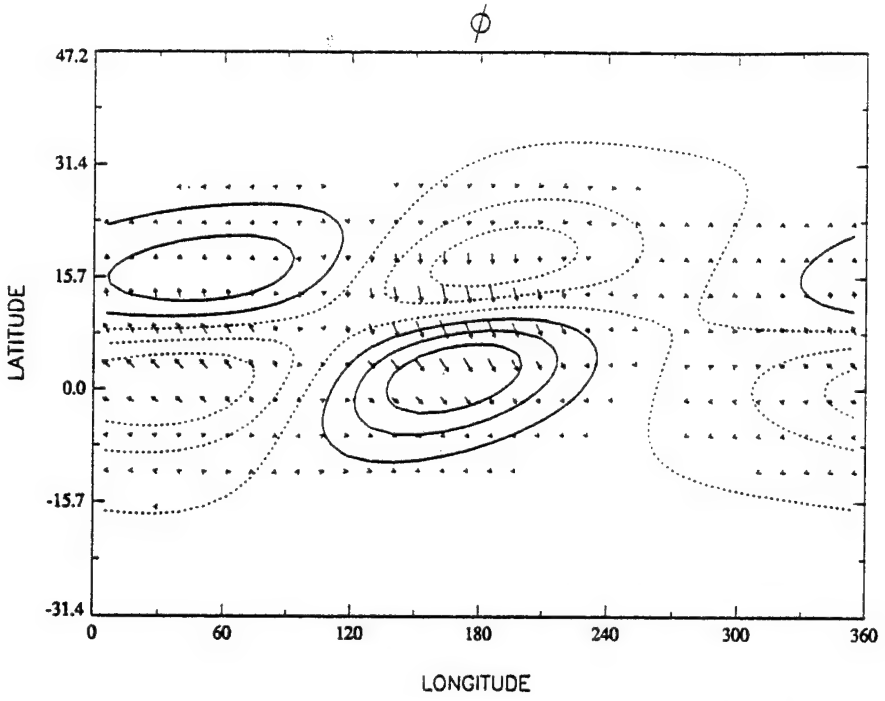


FIG. 8. Structure of global instability for parameters of Fig. 4 but with $\psi = 0.2$.

Dunkerton (1981, 1983) argued that vertical diffusion selects finite $|m|$. With diffusion there is a net growth rate $\omega_i^{\text{eff}} = \omega_i - \nu\epsilon N^2$. In the real world it is uncertain whether diffusion causes scale selection or is a posteriori caused by instability. (In the mesosphere, breaking internal gravity waves will help stabilize the flow.) Other factors may be important. 1) Stability depends on basic-state transience; instabilities must grow rapidly compared to mean flow change. In the tropical mesosphere, planetary waves will modify the basic state on a time scale of 1–5 days. (A Lagrangian time scale is what matters inasmuch as the basic state conserves potential vorticity on fluid parcels and inertial instability is to some extent a parcel instability dependent on anomalous PV.) It would therefore be consistent to require $\gamma \geq 1$–$2\ (\times 10^{-5}\ \text{s}^{-1})$ for instability (cf. Hitchman and Leovy 1986; Hitchman et al. 1987). 2) Realistic basic states contain a continuous spectrum of horizontal and vertical motions. The initial evolution need not be dominated by the most unstable mode; the instability at finite amplitude need not be modal. 3) The meridional circulation itself has a weak stabilizing influence at large ϵ (section 3f).

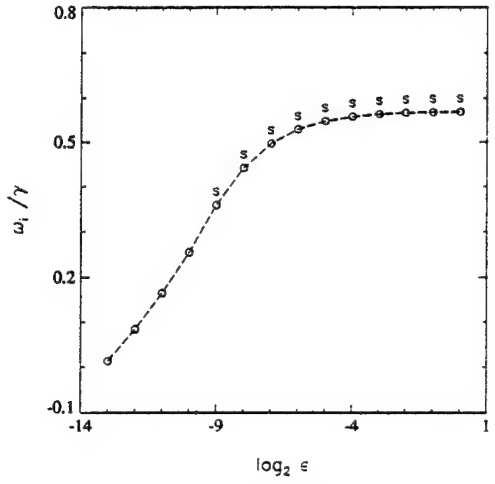


FIG. 9. Growth rate as a function of ϵ.

f. *Effect of meridional circulation*

Cross-equatorial shear in the middle atmosphere is due to diabatic advection by a mean meridional circulation. Including in (2.3) a zonally symmetric component $\bar{v}_0$, independent of latitude, had a minor effect on local stationary instability as shown in Fig. 10. For this series $\gamma = 4 \cdot 10^{-5}\ \text{s}^{-1}$, $\epsilon^{-1} = 256\ \text{m}^2\ \text{s}^{-2}$, $y_c = 1.8 y_s$, $\psi = 1$, and $\bar{u}_0 = 0$. The location of instability was shifted north and east when $\bar{v}_0 > 0$ (not shown). Growth rate was slightly faster in most cases with $\epsilon^{-1} \geq 256\ \text{m}^2\ \text{s}^{-2}$ plotted in Figs. 3, 5, 7, and 9 when $\bar{v}_0 = 5\ \text{m s}^{-1}$. At large ϵ there was a reduction of growth suggesting the possibility of weak scale selection when $\bar{v}_0 > 0$. The physical explanation may be that $\bar{v}_0$ advects the instability away from maximum R, the effect being more dramatic for high-ϵ modes with short meridional scale.

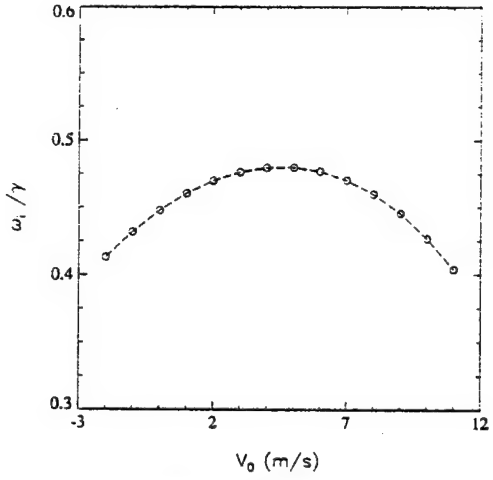


FIG. 10. Growth rate as a function of $\bar{v}_0$.

g. *Results for weak shear*

The preceding results raise two questions about instability in nonparallel flow. 1) Observe from Figs. 4 and 6 that local instability had the same zonal scale as basic-state variation. There was at most one stationary or two propagating cells in the x direction. In general, we expect multicell nonsymmetric instability in locally parallel flow. Under what circumstances, then, can a wave train of many cells develop in the unstable part of the flow? 2) Advected instability was observed at large $|\bar{u}_0|$ because the flow was unstable at all x and periodic boundary conditions were assumed. Real flows could be stable at some longitude. Is advected instability possible in this case? The first question will be addressed now and the second in the next subsection.

It was seen in parallel flow that the minimum zonal scale for latitudinal trapping of eigenfunctions (decay at large $|y|$) is defined by a short-wave cutoff $k_{max}\gamma = \beta$. The argument could be extended to nonparallel flow; minimum zonal scale depends on shear. In the case illustrated above, it was impossible to accomodate several cells in the x direction. For multiple cells γ must be reduced. [That is, in the configuration (3.1). There may exist situations where γ is unimportant outside the region of anomalous PV, such that trapping of eigenfunctions is ensured by the basic state alone, apart from k.]

Figure 11 shows an example in which $\gamma = 0.5 \times 10^{-5}\ s^{-1}$, $\epsilon = 4\ m^{-2}\ s^2$, $y_c = 1.8 y_s$, $\psi = 1$, and $\bar{u}_0 = 0$. Numerical integration used 32 harmonics (note change of axes); results are shown at 240 days. For this value of γ a long time was required to obtain the mode, due to the assumed γ and wave 1 initial condition. (Less time would have been required from high-k initial conditions.) This example demonstrated that multicell instability is possible when the scale of mean-flow variation is much greater than $k_{max}^{-1} \equiv \gamma/\beta$. Another example of small γ is discussed in the Appendix.

What about the opposite situation, when the unstable region is smaller than k_{max}^{-1}? Profiles of the form (3.1)

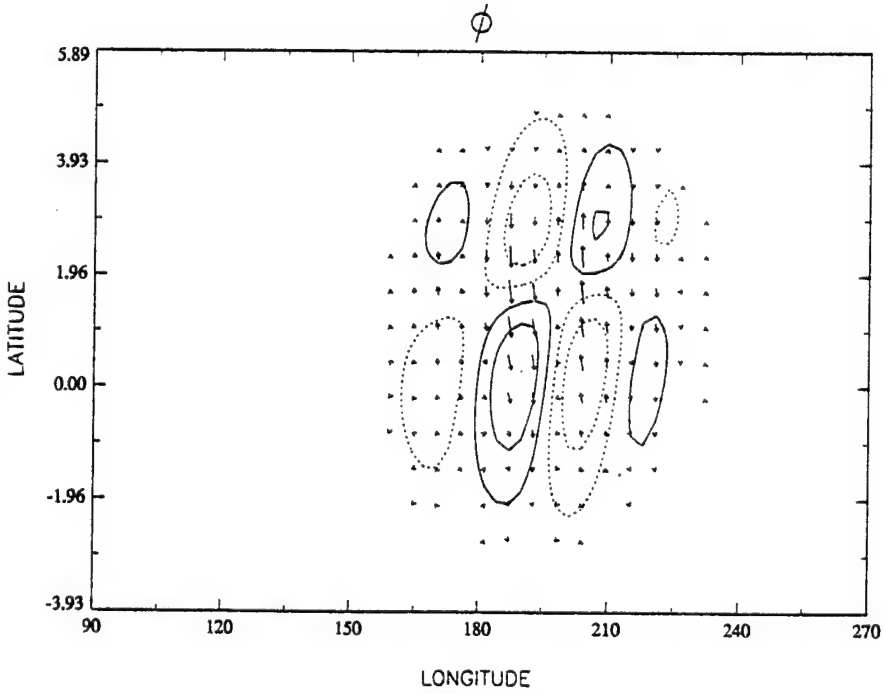


FIG. 11. Structure of multicell instability in weak shear $\gamma = 0.5 \times 10^{-5}\ s^{-1}$.

could be generalized in order to confine the most unstable flow:

$$-\cos(x/a) = 2\sin^2(x/2a) - 1 \rightarrow 2\sin^n(x/2a) - 1. \quad (3.3)$$

Setting $n = 16$, for example, the region of most unstable flow was less than 60° wide. Results indicated that, for parameters of Fig. 4, the local stationary mode was essentially unchanged by reducing the size of the unstable region, although growth rate was reduced slightly (by about 25% when $n = 16$). (An explicit version of the time-dependent code was used to obtain this result and to validate other results of the semi-implicit model.) This supports its interpretation as a local mode that depends, not on global average stability, but on local stability within the most unstable part of the domain (Pierrehumbert 1984).

To be sure, the concept of "local stability" is imprecise when (as in this case) instability and basic state share the same zonal scale. WKB analysis like that of Pierrehumbert (1984) is formally valid only in slowly varying mean flow. Perhaps a more accurate measure of stability would be obtained by averaging, say, over a half-wavelength of perturbation.

h. Sponge region in x

As mentioned in section 2c, the advected instability was altered by a sponge region designed to completely absorb disturbances crossing the domain boundary near $x = 0$. Though artificial, the sponge may simulate regions of stable flow (as undoubtedly exist in the atmosphere) or the tendency of instabilities to "break" and dissipate through nonlinear saturation before making a complete circuit of the globe.

Figure 12 shows growth rate as in Fig. 5 but with sponge region included. Instabilities were unaffected at small $|\bar{u}_0|$ but were stabilized at large $|\bar{u}_0|$. The interpretation is that $\bar{u}_0$ displaced the center of mode downstream from the source of instability near $x = 180°$. For sufficiently large $|\bar{u}_0|$ the mode was pushed into the sponge region and could not grow. Evidently global modes were stabilized by the sponge region also.

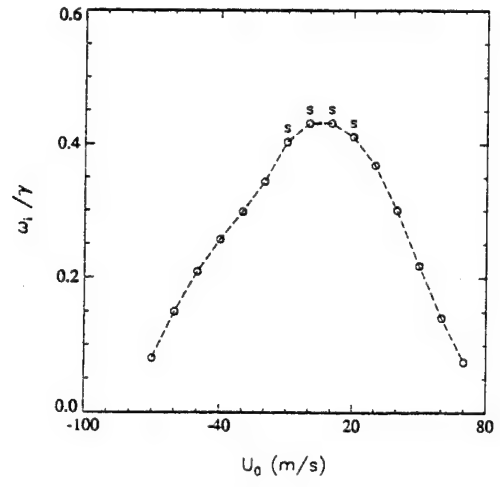


FIG. 12. Growth rate as a function of $\bar{u}_0$ as in Fig. 5, but with sponge region included.

4. Discussion

Results of section 3 documented local as well as global modes of inertial instability in nonparallel flow. Local modes dominate inhomogeneous flow and are likely more important in observed basic states. Local stationary instabilities had the largest growth rate, depending primarily on the local rather than global stability of the basic state: they occurred when (among other things) mean flow advection was small in the unstable region. Large advection caused zonal phase propagation, here referred to as advected instability. Like stationary modes, advected modes had zero group velocity but the center of the wave packet was displaced downstream from the most unstable longitude, and the growth rate reduced. If $\bar{u}_0$ was sufficiently large and the flow stable at large $|x|$, advected modes were stabilized.

Even for a simple profile like (3.1) the parameter dependence of inertial instability is complicated; examples of section 3 highlighted behavior along a few trajectories in parameter space. One of the more interesting results was the dependence of local instability on $\bar{u}_0$, illustrating the breakdown of Galilean invariance in nonparallel flow. Zonal inhomogeneity breaks the symmetry of uniform parallel flow and introduces a new constant: the zonal propagation speed of basic-state pattern. This was zero in (3.1); consequently, instabilities with zero group velocity remained in the same location relative to the mean flow pattern at all times. (Not surprisingly, their growth maximized when the mode best overlapped the most unstable region.) In WKB theory, a flow is said to be *absolutely unstable* when such instabilities exist, that is,

$$\omega_i(k_0) > 0 \quad (4.1)$$

where

$$\left.\frac{\partial\omega}{\partial k}\right|_{k_0} \equiv 0 \quad (4.2)$$

and unstable branches originate on opposite sides of the real-k axis (e.g., Pierrehumbert 1984, 1986). The importance of absolute instability derives from symmetry breaking, for example, in nonparallel/nonuniform flow, or when there is a "wave maker" at some x (Huerre and Monkewitz 1990).

Pierrehumbert (1984) analyzed the two-layer quasi-geostrophic model for baroclinic instability of nonparallel flow, including a discussion of "absolute" and "convective" instability in terms of local and global modes, supplemented by WKB analysis. It is unnecessary to duplicate this work, but simple to quote three of his main conclusions pertaining to local modes:

Growth rate of a local mode is determined by the maximum baroclinicity in the domain, and not by the average baroclinicity.

This was suggested for inertially unstable flow (substituting "inertial instability" for "baroclinicity") by results of Figs. 5 and 12 demonstrating the insensitivity of local stationary instability to a sponge region near $x = 0$, and by discussion at the end of section 3g in which growth rate was only weakly dependent on the size of the unstable region. As in Pierrehumbert (1984), the maximum baroclinicity or maximum inertial instability of the basic state must take into account factors such as boundary conditions, pressure gradients, and $\bar{u}_0$ in addition to the parcel growth rate (e.g., R). On this account we are reminded that equatorial inertial instability is not a pure parcel instability (Dunkerton 1983).

Vanishing absolute growth rate at infinity is not necessary for localization.

This was apparent from examples based on the profile (3.1), which was unstable at all x but displayed a preference for local instabilities in most circumstances.

The contrast between maximum and minimum baroclinicity in the flow determines the extent of localization of the eigenmodes, with high contrast favoring localization.

In a similar way (although it was not discussed extensively above), the zonal scale of local instability increased as ψ was decreased below 1, before the onset of global instability.

Absolute instability does not depend on the validity of WKB in (4.1, 2) since, by definition, the instability must grow indefinitely at a fixed point in x. This property was demonstrated numerically, but not analytically, for local stationary instability. WKB analysis of inertial instability on an equatorial β plane is difficult, although suggestive of absolute instability as discussed in the Appendix. Basic states like Fig. 2 disallow a clean separation of wave and mean-flow scales, and a numerical demonstration of instability is preferable in this case (cf. Jones and Thorpe 1992).

5. Conclusions

A simple model has shown that inertial instabilities in zonally nonuniform cross-equatorial shear take the form of local and global modes. Local modes may be exactly stationary or display zonal phase propagation depending on, among other things, advection by the basic state. Both subtypes of local instability are distinguished from global instability by their zero group velocity and concentration of amplitude within, or downstream from, the region of most unstable flow. These properties, including breakdown of Galilean invariance, are reminiscent of absolute instability—although, apart from the Appendix, none of our results depended on a WKB approximation. Exact stationarity of phase over a finite range of parameters, and measurable dependence of growth rate on the size of unstable region, are possibly attributable to a non-WKB effect.

It was suggested that for equatorial Rossby wave critical layers in the mesosphere, local stationary modes are important given the strength of background shear γ and weakness of in situ advecting current near the zero-wind line. Such a simplification is desirable in this context to avoid the complexity of global and advected modes. Existence of local stationary modes locked to PV anomalies is consistent with middle atmosphere simulations of inertial instability and Rossby wave breaking by O'Sullivan and Hitchman (1992). As a follow-up, it will be worthwhile to show the evolution of unstable modes in an authentic Rossby wave critical layer, and their effect on potential vorticity evolution. This will be done in the sequel.

Local and global modes of inertial instability share many of the same properties: for example, dependence of growth rate on vertical wavelength and downgradient momentum flux. Vertical scale selection may be achieved with diffusion as argued previously, or, as the new analysis suggests, through mean meridional advection. Nonmodal instability (Farrell 1982) on a transient basic state will allow scale selection at finite amplitude. It should be obvious from our results that when the time scale of basic-state variation is much less than γ^{-1}, exponential modes of inertial instability are irrelevant.

Inertial instability can be expected to have three consequences in the tropical middle atmosphere: horizontal redistribution of angular momentum (and potential vorticity), dissipation of laterally propagating Rossby waves, and vertical mixing of constituents. Because of their similarity, local and global modes may contribute alike to mean flow effects, such as inertial adjustment (Dunkerton 1981) and mixing. [The efficiency of inertial adjustment is open to question. Our recent numerical results in the middle atmosphere demonstrate that an adjustment occurs but is incomplete. See also Thorpe and Rotunno (1989) for further discussion in the context of frontal CSI.] Local modes will be more important for Rossby wave dissipation. According to Killworth and McIntyre (1985), barotropic processes are insufficient for time-averaged critical layer absorption as $t \to \infty$.

Acknowledgments. Comments by Frank Crum, Donal O'Sullivan, and two anonymous reviewers were helpful in clarifying the presentation. Some preliminary results of this study were reported by the author in a workshop at the University of Washington, June 1986. This research was supported by the Air Force Office of Scientific Research, Contracts F49620-89-C-0051 and F49620-92-C-0033, and by the National Science Foundation, Grants ATM-8819582 and ATM-9123797.

APPENDIX

Dispersion Relation at Complex k

Analysis of absolute instability in multidimensional flow is difficult: a formal proof of absolute equatorial inertial instability using WKB theory has not been given and will not be attempted here. (I doubt that such a proof is possible for equatorial instability without approximation to the governing equations.) It is interesting nevertheless to calculate eigenfrequencies of (2.10) at complex k. Several limitations of this approach should be noted. 1) Analysis begins with the complex dispersion relation defined numerically by (2.10) with $\mathcal{D} = 0$ rather than the Laplace transform of the time-dependent system; this excludes continuum modes that may be important in reality. 2) The exact dispersion relation is determined numerically over a finite range of k, rather than analytically over all k. The cubic approximation of section 2c (and a quadratic version of it) gave similar results, but their validity could not be guaranteed for all combinations ϵ, k. 3) The infinite β plane could not be simulated; results were obtained in a channel. Channel-dependent modes occur outside the short-wave cutoff (defined by a hyperbola in the complex-k plane extending outward from k_{max} on the real axis). This region will be ignored. 4) Construction of WKB solutions in some cases may require matching across a WKB breakdown point (Pierrehumbert 1984); the full construction will not be attempted but it will be assumed that matching is possible.

To simplify matters, the shear was assumed independent of latitude:

$$\tilde{U} = \gamma(y - y_c)\left[\frac{1 - \psi \cos x/a}{1 + \psi}\right] \equiv \gamma(y - y_c)\Psi_f(x)$$

and $\tilde{V} \equiv 0$, corresponding as closely as possible to the theoretical results of section 2c. Local stationary instability exists near $y_c = y_s$; discussion will center around this choice of y_c. The eigenproblem (2.10) was solved at $x/a = \pi$, the location of maximum shear (where $\Psi_f = 1$), allowing k to be complex. Eigenfrequencies were obtained by a shooting method for several combinations γ, ϵ. Results shown in Fig. 13 were representative when $\epsilon \sim \epsilon_{neut}$ (in this case, $\gamma = 1.5 \times 10^{-5}\ s^{-1}$, $\epsilon = 0.1657\ m^{-2}\ s^2$). Growth rate increased in both directions away from the real axis, and there was a saddle point near $s = ka = (5.55, -1.50)$—not far from the most unstable wavenumber on the real axis. Existence and location of the saddle are a function of parameters (γ, ϵ, Ψ_f, y_c, etc.) so there is nothing special about this example. In the inversion of Laplace transform, the saddle is considered unavoidable when integrating over k_r if unstable branches originate on opposite sides of the real-k axis (Pierrehumbert 1986). This was apparently the case based on numerical results, but an analytical proof is beyond the scope of this paper.

Not all saddles are unavoidable in this problem, particularly those on the imaginary axis resulting from merger of unstable branches originating in the upper half-plane. This sort of behavior was observed when $\bar{u} = \gamma y$.

Knowledge of $\hat{\omega}(k)$ is sufficient to determine new eigenfrequencies $\omega(k)$ from a complex Doppler shift

$$\omega(k) = \hat{\omega}(k) + k\bar{u}_0$$

where $\hat{\omega}$ is given by the complex dispersion relation in the control case $y_c = y_s$ (e.g., Fig. 13). Other effects could be added, such as x diffusion. In WKB theory, breakdown of Galilean invariance is evidently due to the imaginary part of k, as can be seen by translating a wave packet of exponential shape to the left or right. For constant shear, variation of y_c introduces a constant mean flow change $\bar{u}_0 = -\gamma\delta y_c \Psi_f$. The location of saddle point and its associated eigenfrequency change as a function of y_c; this is illustrated in Fig. 14 where comparison is made to time-dependent model results obtained with $\psi = 0.25$. The agreement is rather good. Growth maximized when the mean flow was zero near the center of the unstable region. Due to noticeable asymmetry of $\hat{\omega}(k)$ between upper and lower half-planes, results were not exactly symmetric about $y_c/y_s = 1.0$.

Growth rates were slightly smaller in the time-dependent model, by a few percent—possibly due to y diffusion, included in the time-dependent model but not in the eigenproblem (2.10). (Horizontal diffusivity in the time-dependent model was approximately 2.1

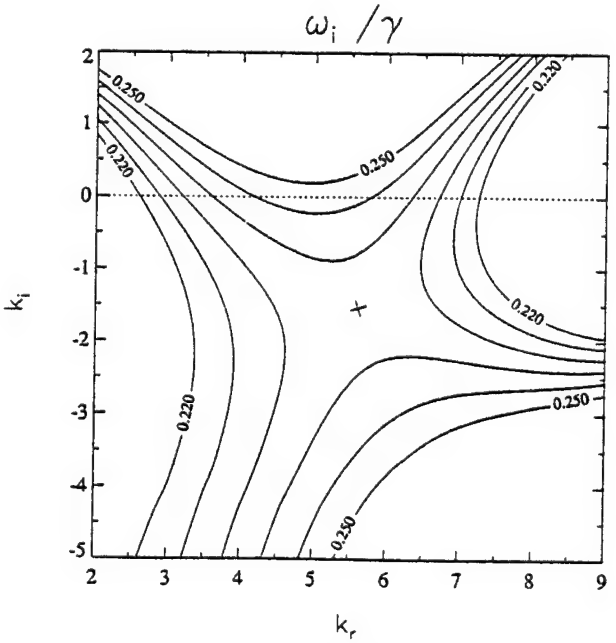


FIG. 13. Inviscid growth rate as a function of complex k obtained from the eigenproblem (2.10) with $\mathcal{D} = 0$. $\bar{u} = \gamma(y - y_c)$, $\gamma = 1.5 \times 10^{-5}\ s^{-1}$, $\epsilon = 0.1657\ m^{-2}\ s^2$, and $y_c = y_s$. (For clarity, not all contours are shown.)

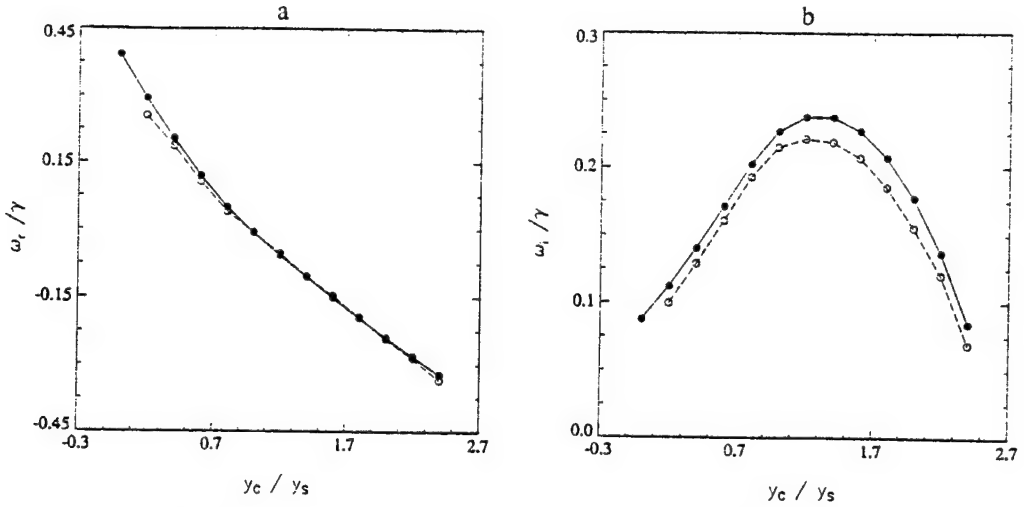


FIG. 14. Frequency (a) and growth rate (b) of local instabilities simulated in time-dependent model (dashed) compared with saddle-point values (solid).

$\times 10^5$ m^2 s^{-1}. Saddle-point values in Fig. 14 were corrected for x diffusion only.) There was also a non-WKB effect such that growth rates in the time-dependent model diminished as ψ increased, that is, as the size of the unstable region was reduced in x. Simulations with $\psi \geq 0.25$ underestimated the true growth rate somewhat.

Calculated and observed periods were in reasonable agreement for advected instabilities. Note that, unlike the situation in Fig. 3, ω_r was nonzero everywhere except near $y_c/y_s = 1.0$; this was due to the smaller value of γ and relatively smaller contribution from the zonally symmetric component of instability.

Displacement of the saddle point off the real axis implies growth of the wave packet between $x/a = \pi$ and its center, which is displaced, in physical space, *downstream* from the region of most unstable flow. This expectation was verified by numerical results, including a more subtle prediction that zonal wavelengths are slightly larger (smaller) for packets displaced to the left (right) of center.

Finally, it was possible to vary x/a, holding other parameters fixed, and determine $k(\omega, x)$ with ω set equal to its saddle-point value as in Fig. 10 of Pierrehumbert (1984). This was a tedious calculation due to rapid variation of k; only two examples were attempted, one at ϵ_{neut} and another above. In both cases, it was striking how quickly k departed from the real axis in both directions as Ψ_f was reduced below unity. This may help explain the locally symmetric appearance of local instability having few cells in the x direction. Also, it may explain why local modes can be found in weakly inhomogeneous flow, as only a slight reduction of Ψ_f below unity immediately moves k into the opposite half-plane—a necessary condition for the existence of local modes (Pierrehumbert 1984).

When $\epsilon \gg \epsilon_{neut}$, the saddle-point trajectory is closer to the imaginary axis, consistent with numerical results demonstrating the preference for quasi-stationary, single-cell instability in this case.

REFERENCES

Andrews, D. G., and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: the generalized Eliassen–Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–2048.

——, J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. Academic Press, 489 pp.

Bennetts, D. A., and B. J. Hoskins, 1979: Conditional symmetric instability—a possible explanation for frontal rainbands. *Quart. J. Roy. Meteor. Soc.*, **105**, 945–962.

Boyd, J. P., 1978: The effects of latitudinal shear on equatorial waves, Part 1: theory and methods. *J. Atmos. Sci.*, **35**, 2236–2258.

——, and Z. D. Christidis, 1982: Low-wavenumber instability on the equatorial beta-plane. *Geophys. Res. Lett.*, **9**, 769–772.

Ciesielski, P. E., D. E. Stevens, R. H. Johnson, and K. R. Dean, 1989: Observational evidence for asymmetric inertial instability. *J. Atmos. Sci.*, **46**, 817–831.

Dunkerton, T. J., 1981: On the inertial stability of the equatorial middle atmosphere. *J. Atmos. Sci.*, **38**, 2354–2364.

——, 1983: A nonsymmetric equatorial inertial instability. *J. Atmos. Sci.*, **40**, 807–813.

Farrell, B. F., 1982: Pulse asymptotics of the Charney baroclinic instability problem. *J. Atmos. Sci.*, **39**, 507–517.

Fritts, D. C., L. Yuan, M. H. Hitchman, L. Coy, E. Kudeki, and R. F. Woodman, 1992: Dynamics of the equatorial mesosphere observed using the Jicamarca MST radar during June and August 1987. *J. Atmos. Sci.*, **49**, 2353–2371.

Held, I. M., and A. Y. Hou, 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.*, **37**, 515–533.

Hitchman, M. H., and C. B. Leovy, 1986: Evolution of the zonal mean state in the equatorial middle atmosphere during October 1978–May 1979. *J. Atmos. Sci.*, **43**, 3159–3176.

——, ——, J. C. Gille, and P. L. Bailey, 1987: Quasi-stationary zonally asymmetric circulations in the equatorial lower mesosphere. *J. Atmos. Sci.*, **44**, 2219–2236.

Holton, J. R., 1975: *The Dynamic Meteorology of the Stratosphere and Mesosphere*. Amer. Meteor. Soc., 319 pp.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential-vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946.

Huerre, P., and P. A. Monkewitz, 1990: Local and global instabilities in spatially developing flows. *Ann. Rev. Fluid Mech.*, **22**, 473–537.

Hunt, B. G., 1981: The maintenance of the zonal mean state of the upper atmosphere as represented in a three-dimensional general circulation model extending to 100 km. *J. Atmos. Sci.*, **38**, 2172–2186.

Jones, S. C., and A. J. Thorpe, 1992: The three-dimensional nature of 'symmetric' instability. *Quart. J. Roy. Meteor. Soc.*, **118**, 227–258.

Killworth, P. D., and M. E. McIntyre, 1985: Do Rossby-wave critical layers absorb, reflect, or over-reflect? *J. Fluid Mech.*, **161**, 449–492.

Krishnamurti, T. N., P. K. Jayakumar, J. Sheng, N. Surgi, and A. Kumar, 1985: Divergent circulations on the 30 to 50 day time scale. *J. Atmos. Sci.*, **42**, 364–375.

——, D. K. Oosterhof, and A. V. Mehta, 1988: Air–sea interaction on the time scale of 30 to 50 days. *J. Atmos. Sci.*, **45**, 1304–1322.

Liebmann, B., 1987: Observed relationships between large-scale tropical convection and the tropical circulation on subseasonal time scales during the Northern Hemisphere winter. *J. Atmos. Sci.*, **44**, 2543–2561.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan*, **44**, 25–43.

O'Sullivan, D. J., and M. H. Hitchman, 1992: Inertial instability and Rossby wave breaking in a numerical model. *J. Atmos. Sci.*, **49**, 991–1002.

Philander, G., 1989: *El Niño, La Niña, and the Southern Oscillation.* Academic Press, 293 pp.

Pierrehumbert, R. T., 1984: Local and global baroclinic instability of zonally varying flow. *J. Atmos. Sci.*, **41**, 2141–2162.

——, 1986: Spatially amplifying modes of the Charney baroclinic instability problem. *J. Fluid Mech.*, **170**, 293–317.

Stevens, D. E., 1983: On symmetric stability and instability of zonal mean flows near the equator. *J. Atmos. Sci.*, **40**, 882–893.

——, and P. Ciesielski, 1986: Inertial instability of horizontally sheared flow away from the equator. *J. Atmos. Sci.*, **43**, 2845–2856.

Thorpe, A. J., and R. Rotunno, 1989: Nonlinear aspects of symmetric instability. *J. Atmos. Sci.*, **46**, 1285–1299.

Seasonal Development of the Extratropical QBO in a Numerical Model of the Middle Atmosphere

Donal O'Sullivan and Timothy J. Dunkerton

Reprinted from JOURNAL OF THE ATMOSPHERIC SCIENCES, Vol. 51, No. 24, 15 December 1994
American Meteorological Society

Seasonal Development of the Extratropical QBO in a Numerical Model of the Middle Atmosphere

DONAL O'SULLIVAN AND TIMOTHY J. DUNKERTON

Northwest Research Associates, Bellevue, Washington

(Manuscript received 8 November 1993, in final form 4 May 1994)

ABSTRACT

The seasonal (wintertime) development of middle atmosphere circulation in opposite phases of the equatorial quasi-biennial oscillation (QBO) was simulated with a three-dimensional nonlinear numerical model. In the stratosphere, the effect of equatorial QBO was generally consistent with the extratropical QBO observed by Holton and Tan, namely, a stronger midwinter polar vortex in the westerly phase, and vice versa. However, the extratropical response to the QBO was sensitive to other factors such as mesospheric gravity wave drag and the amplitude of Rossby waves specified at the model's lower boundary. The extratropical QBO was realistic only when a drag parameterization was included and Rossby wave amplitudes lay in an intermediate range close to the observed. At somewhat stronger forcing, the model's response was largest in the mesosphere where (in this case) westerlies were stronger in the easterly phase of equatorial QBO. This was apparently due to a shielding effect.

The theory of planetary wave–mean flow interaction suggests that the sensitivity to equatorial QBO should be greatest for wave forcings near a "bifurcation" point. Below this threshold the stratosphere approaches radiative equilibrium, shutting off vertical propagation of planetary waves. Supercritical forcing leads to a major warming. The model's sensitivity to forcing, while consistent with this idea, was most apparent in perpetual solstice runs without parameterized wave drag. Seasonal integrations with wave drag produced a more realistic extratropical QBO, making the bifurcation less conspicuous.

1. Introduction

It is currently thought that the equatorial quasi-biennial oscillation (QBO) affects the wintertime circulation of the extratropical middle atmosphere. The stratospheric polar vortex is weaker and more disturbed by planetary waves in the easterly phase of the QBO (Holton and Tan 1980, 1982; Wallace and Chang 1982; Labitzke 1982, 1987; van Loon and Labitzke 1987; Dunkerton and Baldwin 1991; Baldwin and Dunkerton 1991). During the westerly phase there is a colder, stronger vortex with greater potential for ozone depletion (Bojkov 1986; Garcia and Solomon 1987; Lait et al. 1989). In the Northern Hemisphere, this "extratropical QBO" is strongest in early to midwinter (Dunkerton and Baldwin 1991, 1992). It is not uniform in all years but (for reasons not well understood) seems more prominent near solar minima (Labitzke and van Loon 1988) in late winter (Dunkerton and Baldwin 1992).

Planetary waves play an important role in the extratropical QBO as measured by their Eliassen–Palm flux and divergence (Dunkerton and Baldwin 1991—their Fig. 16). Whether they are the main cause of the extratropical QBO is not readily answered from observations, but is suggested by recent numerical investigations (O'Sullivan and Salby 1990; Dameris and Ebel 1990; Holton and Austin 1991; O'Sullivan and Young 1992). Starting from initial conditions differing only in the phase of equatorial QBO (with planetary Rossby waves forced at a tropopause lower boundary), twin experiments gradually diverge over several weeks until the extratropical mean flow is weakened by wave transport—the overall deceleration being somewhat larger in the easterly phase of the QBO. There can be a significant extratropical difference induced by the equatorial QBO that is most realistic for intermediate wave forcings (Holton and Austin 1991; O'Sullivan and Young 1992). Large forcings cause the vortex to break down prematurely (as in a major warming) irrespective of the QBO, while weak forcings produce a weak or negligible effect.

According to numerical results, it is not the instantaneous *linear* difference of planetary wave structure due to equatorial QBO that is important, but rather the cumulative effect of planetary wave–mean flow interaction over a period of time.[1] The model's extratropical

Corresponding author address: Dr. Timothy J. Dunkerton, Northwest Research Associates, P.O. Box 3027, Bellevue, WA 98009-3027.

[1] Linear structural changes are possible, but their effect on EP flux divergence is confined mainly near subtropical critical latitudes (O'Sullivan and Young 1992).

© 1994 American Meteorological Society

QBO is therefore a quasi-linear or nonlinear phenomenon rather than a linear one.

By the same token, the extratropical QBO evolves during winter. The mean zonal wind difference between west and east phases is maximum in midwinter (December–January), while that of the EP flux maximizes a bit earlier (November–December) (Dunkerton and Baldwin 1991—their Figs. 10, 20). To model the seasonal development realistically requires time-dependent forcings and diabatic cooling of the observed magnitude. O'Sullivan and Young (1992) ignored the seasonal development and relaxed the flow to an equilibrium zonal wind similar to the observed (50% stronger than their initial condition in midlatitudes), rather than radiative equilibrium. Consequently their radiative forcing was too weak, as was the induced meridional circulation. Holton and Austin (1991) used a sophisticated radiative algorithm, but began their experiments in mid-January, which is after the maximum of observed extratropical QBO.

One purpose of this study was to model the seasonal development of extratropical QBO with realistic thermal drive; another was to understand its sensitivity to various parameters, such as the amplitude of Rossby waves entering the stratosphere. Theory suggests that weak waves allow the mean state to approach radiative equilibrium, impeding further wave propagation. Conversely, strong waves decelerate the mean flow, enhancing propagation until the vortex experiences a major warming. Simple one-dimensional models predict a "bifurcation" between extremes that is a function of wave forcing amplitude and other parameters (Holton and Mass 1976; Holton and Dunkerton 1978; Yoden 1987a,b,c, 1990).

In light of the 1D theory, it is interesting to consider the effect of some external influence, like the equatorial QBO, in conjunction with planetary waves forced at the extratropical tropopause. If the wave forcing is such that for one phase of the QBO the forcing is subcritical, while for the opposite phase it is supercritical, we expect maximum sensitivity to the equatorial QBO. If the forcing is above or below this threshold in both phases alike, we expect minimum sensitivity. This behavior is consistent with previous numerical studies of the extratropical QBO but deserves further investigation, for example, to determine whether the bifurcation occurs in a nonlinear, three-dimensional model of the middle atmosphere,[2] and what effect the seasonal cycle has. Other factors may also affect the extratropical response to the QBO. The effect of gravity wave drag on planetary wave–mean flow interaction will be emphasized in the discussion to follow.

[2] Holton and Wehrbein (1981) demonstrated that vacillation was possible in a 3D quasi-linear model, and nonlinear barotropic integrations showed evidence of bifurcation (Salby et al. 1990; Yoden and Ishioka 1993).

Section 2 briefly describes the numerical model and the parameterization of thermal and mechanical forcings. Section 3 describes the seasonal mean-flow evolution with and without parameterized wave drag, and the variation of extratropical QBO as a function of wave forcing in a baseline series of experiments. These results are discussed further in section 4 along with possible causes of extratropical QBO variability.

2. Numerical model

The numerical model used in this study was the same as that of O'Sullivan and Young (1992): a global, primitive equation model extending from 10 to 70 km with geopotential specified at the lower boundary. Simulations were performed at T42/21 horizontal resolution (spherical harmonics with total wavenumber up to 42 and zonal wavenumber up to 21 are included) with 20 equally spaced levels in log p ($\Delta z = 3$ km). Sixth-order horizontal diffusion was included to suppress grid-scale motions ($\nu = 2.5 \times 10^{28}$ m^6 s^{-1}). The time step was 15 minutes. Time integrations began on "11 September" and extended thereafter for 200 days. The mean thermal forcing and perturbation lower-boundary condition varied through the seasonal cycle as described below.

The main differences relative to O'Sullivan and Young (1992) were the parameterization of thermal and mechanical drive, modified to produce a more realistic wintertime circulation, and integration through a seasonal cycle. Radiative heating and unresolved body forces were parameterized as linear relaxation. This approach allowed many integrations at relatively low cost.

a. *Thermal and mechanical forcings*

Zonally averaged diabatic heating was of the form

$$\bar{Q}(y, z, t) = \alpha_T(z)[\bar{T}_E(y, z, t) - \bar{T}(y, z, t)], \quad (2.1)$$

where $\bar{T}_E$ is the "radiatively determined" profile of Shine (1987). This is not the true radiative equilibrium under perpetual conditions but a radiatively determined temperature in the seasonal cycle. When used in conjunction with a temporally fixed Newtonian cooling rate,

$$\alpha_T(z) = \left[0.09 + 0.04 \tanh\left(\frac{z - 35}{10}\right)\right] \text{day}^{-1}, \quad (2.2)$$

realistic solstitial heating and cooling are obtained. This was verified by comparison to K. Shine's MIDRAD algorithm in a 2D model. Newtonian cooling rates implied by that algorithm vary in latitude and time, an effect which is ignored in (2.2) but is implicit in our use of radiatively determined, rather than radiative equilibrium, temperature. Identical damping was applied to the deviation from zonal-mean temperature.

The temporal dependence of radiative forcing was further simplified by adopting only the solstitial profile of $\bar{T}_E$, holding its symmetric part about the equator fixed while letting the antisymmetric part vary sinusoidally through the annual cycle.

As noted by Shine (1987) and many others, the radiatively determined state in the middle atmosphere yields an enormous zonal wind at winter solstice. To overcome this problem requires body forces that retard the mean flow. Two kinds of body force can play a role. (i) Planetary Rossby waves forced at the model's lower boundary (section 2b) decelerate the polar vortex. These waves can have a large effect, but (in the absence of other body forces) their ability to retard the vortex turns out to be quite sensitive to forcing amplitude (section 3a). (ii) Unresolved disturbances such as gravity waves exert a substantial drag on the mean flow, particularly in the mesosphere.

An investigation of gravity waves in the middle atmosphere's *three-dimensional* circulation would be interesting (Dunkerton and Butchart 1984) but outside the scope of this paper. It will be assumed instead that the cumulative effect of unresolved waves is to retard the zonal-mean westerlies to something like their climatological value. This was achieved in the numerical model by inserting in the zonally averaged zonal momentum equation a one-way friction (Dunkerton 1991) designed to relax overly strong westerlies to the climatological average profile of mean zonal wind $\bar{u}_M$ (Barnett and Corney 1985) if and only if $\bar{u} > \bar{u}_M$. Profiles of monthly mean $\bar{u}_M$ were derived from Barnett and Corney for each of the 12 calendar months, and interpolated to the current time step. The mechanical damping coefficient was zero for winds below 25 m s^{-1}, increasing to 0.5 per day for winds in excess of 50 m s^{-1}. One-way friction had little effect on the polar lower stratosphere and no direct effect in the Tropics or summer hemisphere, but severely restricted the mesospheric jet and thereby maintained a strong diabatic circulation at upper levels. Hereafter the friction will be referred to as "parameterized wave drag."

Although this device was used mainly for simplicity, gravity wave momentum flux divergence can be written as linear relaxation when momentum flux is proportional to the negative of mean flow (as suggested by some observations), and the mean flow varies slowly over a density scale height. Similar results might be achieved with a more sophisticated gravity wave parameterization (Fritts and Lu 1993) tuned to produced realistic winds.

b. *Rossby wave forcing*

Geopotential was specified at the lower boundary (10 km). The forcing consisted solely of stationary zonal wavenumber 1 with maximum amplitude at 60°N as in Eq. (2) of O'Sullivan and Young (1992). It was turned on gradually over the first ten days to an initial value 150 m at $t = t_0$ ($\equiv$ September 21). Thereafter, a slow seasonal amplification was assumed

$$\phi_{60N}^2 = 150^2 + b^2 \sin\left[\frac{\pi(t - t_0)}{180 \text{ days}}\right], \quad (2.3)$$

as suggested by Randel's (1992) climatology. Here ϕ_{60N} is the maximum value of wavenumber one amplitude on the lower boundary. The parameter b was varied from one experiment to the next to investigate the effect of forcing amplitude on the extratropical QBO. Hereafter we denote this variation in terms of a parameter ϕ_m, the maximum value of ϕ_{60N} attained at winter solstice. Values of ϕ_m ranged from 150–350 m, like those observed (~200–280 m).

Slowly varying forcing is an idealization of the real atmosphere with its transient waves. Pulsations of wave activity cause episodes of enhanced mixing and deceleration (O'Sullivan and Salby 1990), but so long as quasi-stationary waves dominate the spectrum, the extratropical QBO should be much the same.

c. *Equatorial QBO*

The model was initialized with an equinoctial zonal wind profile typical of September, together with a contribution from the equatorial QBO like that of O'Sullivan and Young (1992)—a Gaussian profile in latitude and height, centered on the equator between about 20 and 40 km—but (in our case) having an e-folding width of 10° (15°) latitude and maximum amplitude 15 m s^{-1} (−20 m s^{-1}) in westerly (easterly) phase. Numerical experiments showed that the extratropical QBO was unrealistically large for wider equatorial QBOs. Interestingly, most of the extratropical QBO was due to the westerly equatorial phase; results in easterly phase did not differ much from control experiments with no QBO.

West-phase experiments included a small restoring force to prevent erosion of equatorial winds. This device might represent the positive acceleration associated with equatorial Kelvin waves not included in the model.

3. Results

a. *Experiments without wave drag parameterization*

O'Sullivan and Young (1992) relaxed mean temperature to an equilibrium state similar to the observed. Consequently their mean zonal wind was realistic but diabatic cooling and mean meridional circulation were not. Here mean temperature was relaxed to a radiatively determined state so that, without dynamical heat fluxes or opposing body forces, the polar vortex becomes extremely strong. This "cold bias" can be alleviated by transport due to planetary Rossby waves, gravity waves, or both. It is interesting (and perhaps amusing) to consider first the situation without gravity waves.

where vertically propagating planetary waves alone are responsible for mean flow deceleration. Our largest extratropical QBO was obtained in perpetual solstice runs (temporally constant $\bar{T}_E$, ϕ_m after the initial 10-day ramp) without parameterized wave drag. In some of these, the time-averaged difference of mean zonal wind between west and east QBO phases exceeded 160 m s^{-1} in the mesosphere, the pattern being similar in other respects to that of O'Sullivan and Young (1992; their Fig. 6a).

Time series illustrating the QBO's effect are shown in Fig. 1. The mean zonal wind in both cases accelerated under radiative control until day 50, after which the east, but not west, case decelerated—gradually at first, and then more rapidly. The magnitude of this "extratropical QBO" at days 120–150 (ten times the observed) revealed, in essence, the difference between a radiatively determined state (west phase) and a major warming (east phase). The time-averaged difference was greatest when the averaging interval was located after the first major warming in the east-phase run, and wave forcing was optimally chosen to lie near a "bifurcation point" so there was no major warming in the west-phase run. The bifurcation was anticipated by simple one-dimensional models (Holton and Mass 1976; Holton and Dunkerton 1978) and its effect in this case was quite dramatic. In east-phase experiments (similar to those with no equatorial QBO), wave forcings $\phi_m \geq 300$ m led to a major warming. Smaller forcings (e.g., $\phi_m \leq 275$ m) could not prevent the approach to radiative equilibrium. The effect of the QBO was to shift the bifurcation point toward greater ϕ_m in the west phase, so that twin experiments with identical wave forcing near the original bifurcation produced a large west-minus-east (W–E) difference. Remarkably the bifurcation was just as sharp in the 3D model as in simple 1D calculations of Holton and Dunkerton (1978; see also Yoden 1987a,b).

Comparable, though less dramatic, results were obtained when the seasonal cycle was included (without parameterized wave drag) as in sections 2a,b. Figure 2 shows the time-averaged December–February (DJF) mean zonal flow in a pair of runs with $\phi_m = 375$ m (Figs. 2a,b), and their difference (Fig. 2c). The averaging interval was 90 days, beginning on day 80 of the model run (December 1); this interval will apply throughout the paper unless otherwise noted. None of the experiments revealed any significant difference between east and west phase in early winter (prior to December), even with parameterized wave drag, so there was little point in including this time period in the averaging interval. [The model's early winter was defective, since observations show a QBO effect in November (Dunkerton and Baldwin 1991).] As in perpetual solstice runs near the bifurcation, there was a major warming in east phase (Fig. 2b), but not west phase (Fig. 2a), resulting in a large DJF W–E difference (Fig. 2c). The difference was exaggerated by the

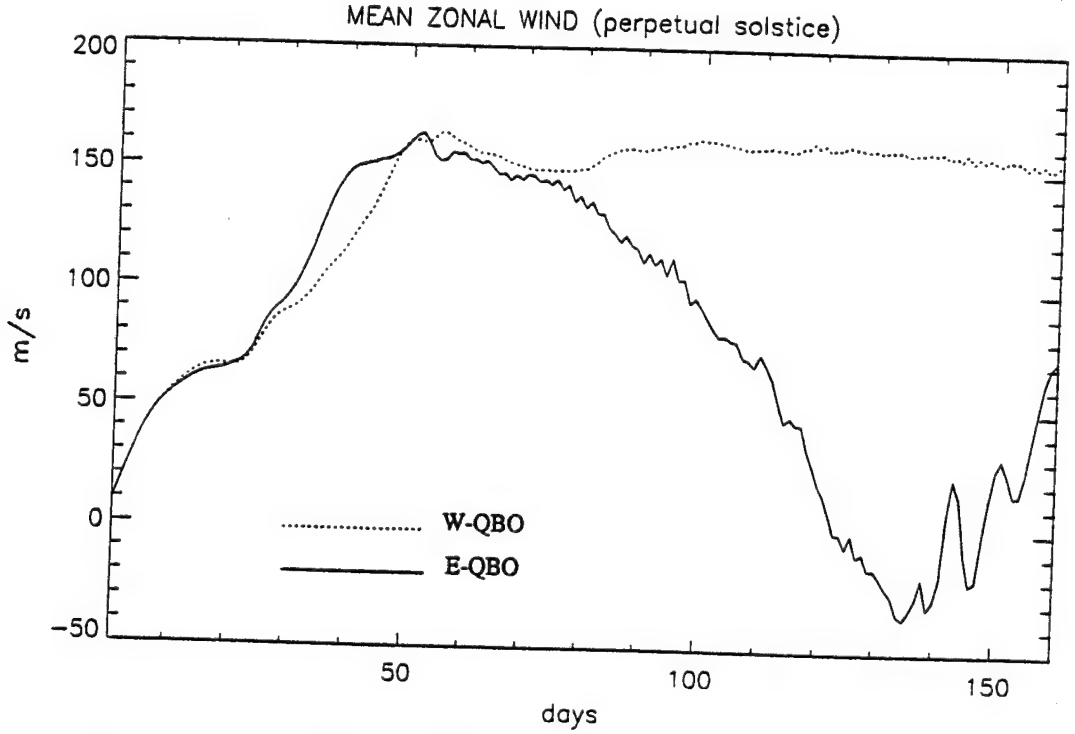


FIG. 1. Time series of mean zonal wind at 75°N, 0.5 mb in perpetual solstice runs without parameterized wave drag; forcing amplitude $\phi_m = 325$ m constant in time.

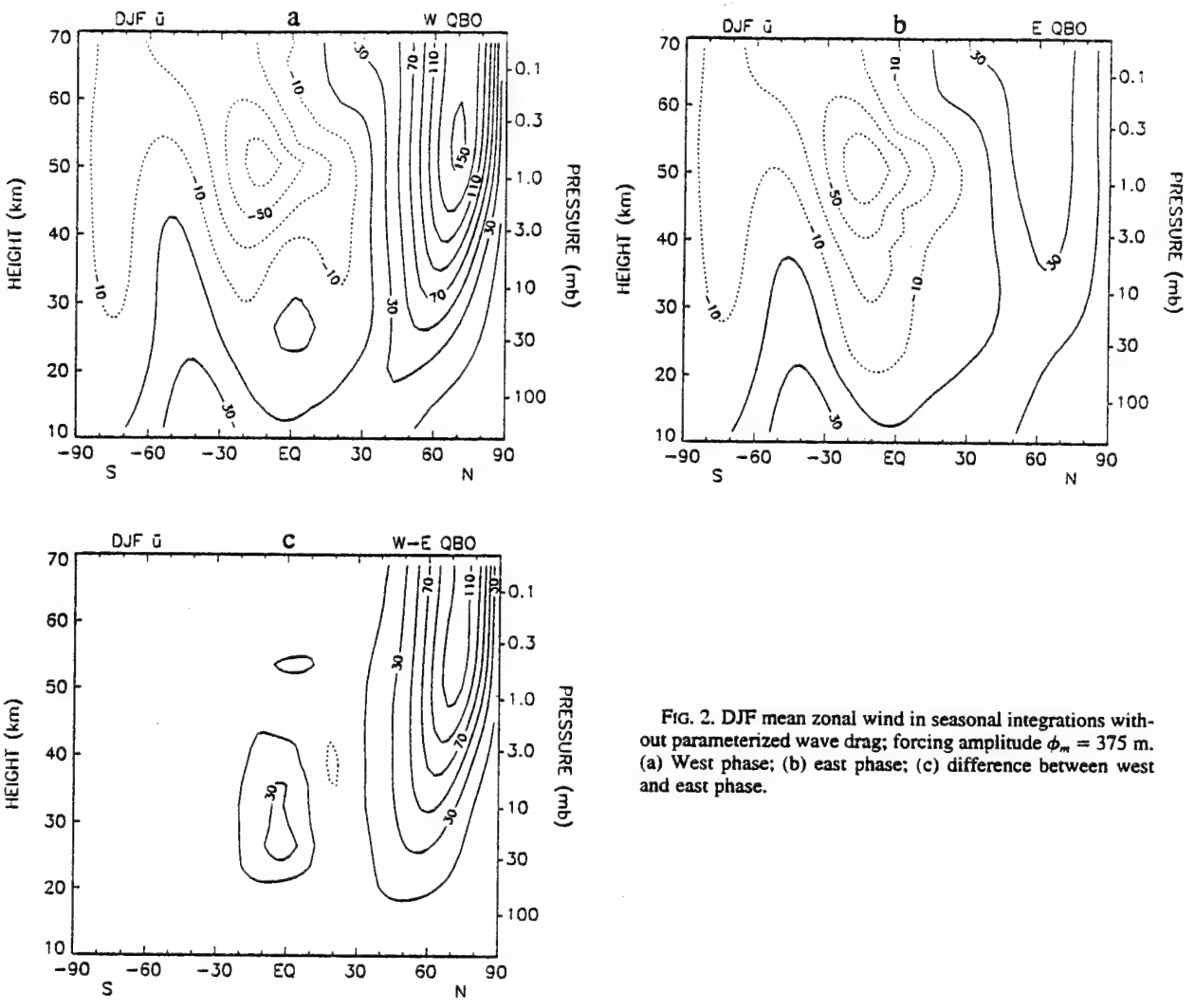


FIG. 2. DJF mean zonal wind in seasonal integrations without parameterized wave drag; forcing amplitude ϕ_m = 375 m. (a) West phase; (b) east phase; (c) difference between west and east phase.

approach to radiatively determined conditions in the west phase, resulting in a mean flow (Fig. 2a) almost the same as if there were no planetary wave forcing at all.

Figures 3a,b display the seasonal evolution of mean zonal wind at 75°N, 0.5 mb in west and east phase, respectively, in several runs without parameterized wave drag. These time series can be compared to Yoden's (1990, his Fig. 5) and are qualitatively similar. West-phase experiments were radiatively determined, except those with forcings in excess of 400 m. Mean zonal wind near the pole increased slightly with increasing forcing, due to asymmetric displacement of the vortex off the pole (which should not obscure the fact that integrated angular momentum is diminished by the vertical EP flux entering the domain at the lower boundary). East-phase experiments were perturbed by forcings exceeding 300 m; some of these contained strong vacillations after the initial breakdown. Although sharp bifurcations were not apparent, there was a maximum sensitivity of DJF mean flow to forcing amplitude, located between 400 and 450 m in the west phase and 325 and 375 m in the east phase. Forcings between 350 and 400 m therefore maximized the effect of the QBO (cf. Fig. 2c).

The QBO's influence could be visualized as a shift of the bifurcation point, which in this case is no longer a "point" (as in perpetual solstice runs) but rather a range of ϕ_m values over which the time-averaged flow is most sensitive to changes in ϕ_m. The shift is toward greater ϕ_m in the west phase, compared to east-phase integrations and control runs with no QBO. Alternatively, the QBO's effect is analogous to a change in ϕ_m, a flip from west to east phase being roughly equivalent to a 75–100 m increase of forcing.

In seasonal integrations, the response of time-averaged flow to Rossby wave forcing was therefore a smooth function of ϕ_m, unlike perpetual solstice runs with a near-discontinuity at the bifurcation point. The radiative drive and lower boundary forcing varied as a

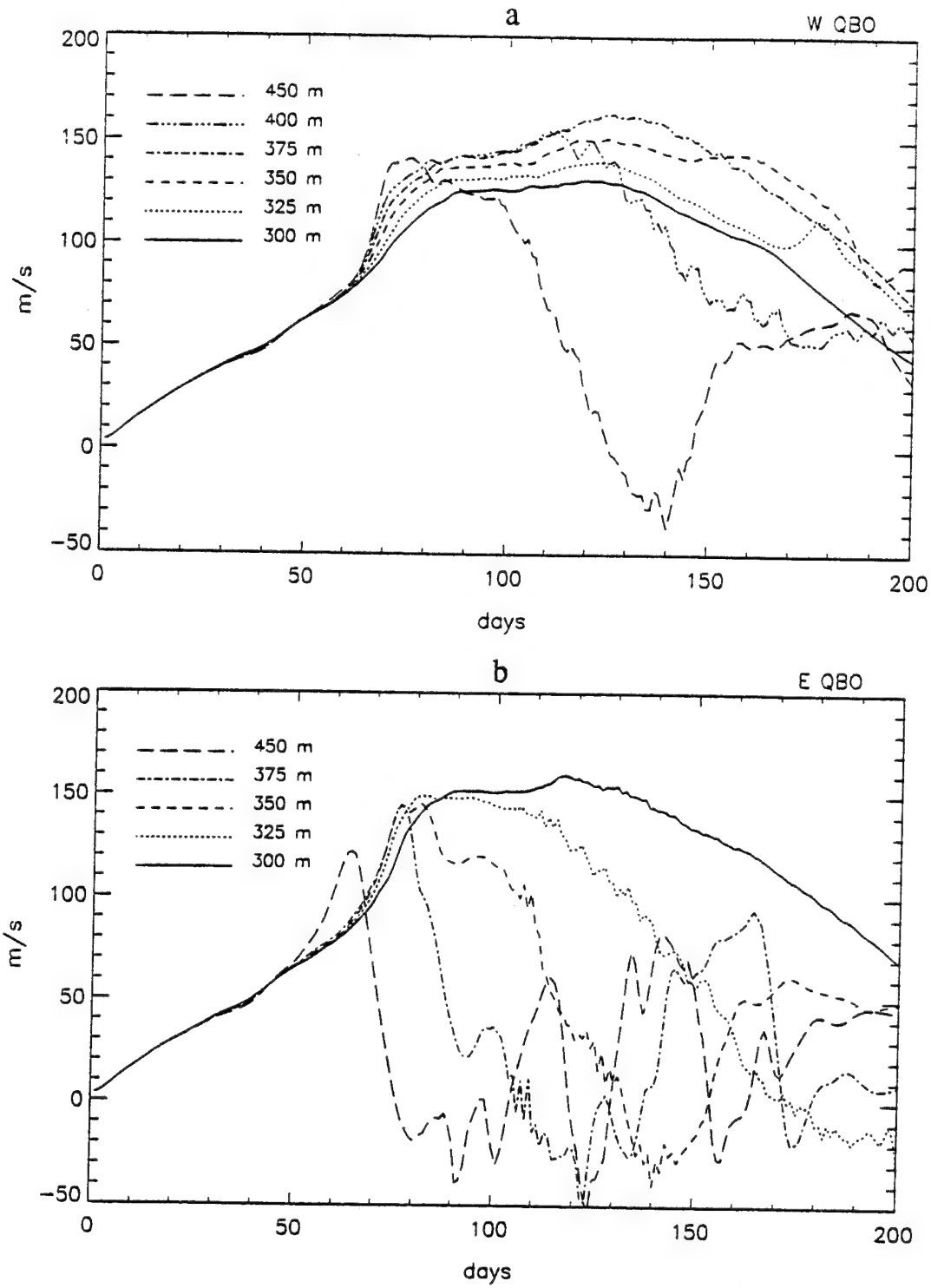


FIG. 3. Time series of mean zonal wind at 75°N, 0.5 mb in seasonal integrations without parameterized wave drag. (a) West phase; (b) east phase.

function of time. It was not merely a question of whether a major warming would occur for a particular ϕ_m (as in perpetual solstice runs), but when such a warming would occur and (if so) how large it would be. In seasonal integrations there are more degrees of freedom, as it were, that determine the response to ϕ_m.

Despite the added complexity, it is useful to imagine a "fuzzy bifurcation" in the response, demarcating the range of ϕ_m values over which the stratospheric flow changes from an essentially radiative state (at small ϕ_m) to a highly disturbed state (at large ϕ_m).

The DJF time period was obviously important to the seasonal development of extratropical QBO in the model and is emphasized for the following reasons. As already noted, differences were insignificant in early winter, prior to December. Nor was there a large QBO effect in late winter, after February. The polar vortex breaks down as spring approaches; differences due to the QBO are not expected once this has happened. Late winter warmings are a consequence of diminished radiative forcing and increasing susceptibility to wave transport as the vortex erodes away; these breakdowns are expected regardless of the phase of the QBO (Holton and Austin 1991). Large W–E differences are obtained when a midwinter warming occurs in one phase but not in the other, and the averaging interval is chosen accordingly, as in Fig. 2. The 90-day DJF averaging interval (days 80–170) captured the W–E difference well, as in observations (Dunkerton and Baldwin 1991, 1992). The modeled seasonal development was fairly realistic but, as already noted, its early winter was too quiescent and devoid of any QBO effect.

Needless to say, the model's extratropical QBO was well defined in midwinter. In the real atmosphere the situation is not so clear cut. While there is a tendency for major midwinter warmings to occur preferentially in the east phase, this is hardly significant, since warmings also occur in the west phase and, in any case, the WMO definition of a "major" warming is arbitrary. Observations indicate a significant, but much smaller, composite difference than shown in Fig. 2c.

Although the observed W–E difference may be due, in part, to interannual variations unrelated to the QBO (e.g., tropospheric forcing), this explanation is inadequate. The simulated difference (Fig. 2c) came from an unrealistic radiatively determined state in the west phase (Fig. 2a), which is not observed in either hemisphere. Obviously important are transport mechanisms, other than vertically propagating planetary waves, that restrict the amplitude of the mesospheric jet. The mesosphere contains barotropic instabilities (Hartmann 1983) that redistribute angular momentum horizontally and homogenize the high-latitude flow (Dunkerton and Delisi 1985a) and large-amplitude gravity waves (Fritts 1984) that transport momentum vertically and are thought to decelerate the flow on the order of 100 m s^{-1} day^{-1} (Lindzen 1981).

While interesting in their own right, these transport mechanisms can also affect planetary wave propagation. Diminished mean flow in high latitudes alters the vertical propagation of planetary waves and affects the bifurcation property of the vortex. A weakened vortex enables vertical Rossby wave propagation (Charney and Drazin 1961), increasing the probability of wave–mean flow interaction and vortex breakdown.

b. *Experiments with wave drag parameterization*

Numerical integrations with parameterized wave drag produced a more realistic polar night jet and, as a result, a more realistic extratropical QBO. Figures 4a,b show latitude–time plots of mean zonal wind (west and east phase, respectively) at 8 mb, with planetary wave forcing ϕ_m = 275 m. Both experiments were "subcritical" in that neither led to a major midwinter warming. There was some contraction and deceleration of the jet after the midpoint of integration, the details differing slightly between the two experiments. At this level, simulations did not diverge significantly until after 160 days, when a breakdown occurred in the east phase. Similar evolution was observed at ϕ_m = 250 m (not shown).

With slightly larger forcing (ϕ_m = 300 m) there was a major midwinter warming in the east (but not west) phase (Figs. 5a,b). As shown later, this pair of integrations gave the maximum W–E difference in the stratosphere, the sign, shape, and magnitude of $\Delta\bar{u}$ being consistent with observations. There was an overall reduction of DJF mean flow north of about 45°N and slight acceleration to the south (note formation of a weak jet near 30°N between days 120 and 140 in Fig. 5b). Similar evolution was observed at ϕ_m = 325 m (not shown) except that a minor midwinter breakdown also occurred in the west phase.

At still higher forcing (ϕ_m = 350 m), midwinter warmings occurred in both QBO phases (Figs. 6a,b), although the patterns were not identical. The east phase event qualified as a "major" sudden warming by day 100, followed by some strong vacillations, while the west phase evolution (after an initial deceleration at day 85) resembled the gradual deceleration and breakdown seen in previous examples. Apparently the west phase transition or "fuzzy bifurcation" lay between 300 and 350 m, whereas the east phase transition was between 275 and 300 m. We emphasize, however, that these transitions (as a function of ϕ_m), like Fig. 3, were nowhere as abrupt as those observed in perpetual solstice runs without parameterized wave drag.

DJF W–E mean zonal wind differences for several experiments (ϕ_m = 275, 300, 325, 350 m) are summarized in Figs. 7a–d. The extratropical QBO began at low forcing with a positive anomaly in the polar mesosphere (Fig. 7a) that descended, with increasing forcing, to the stratosphere (Fig. 7b). Simultaneously a weak negative anomaly appeared in the midlatitude stratosphere, arching into the mesosphere on top of the positive anomaly (Fig. 7c). At large forcing both positive and negative components weakened to insignificant values (Fig. 7d).

This sequence can be contrasted to O'Sullivan and Young (1992), where the basic QBO dipole—a pos-

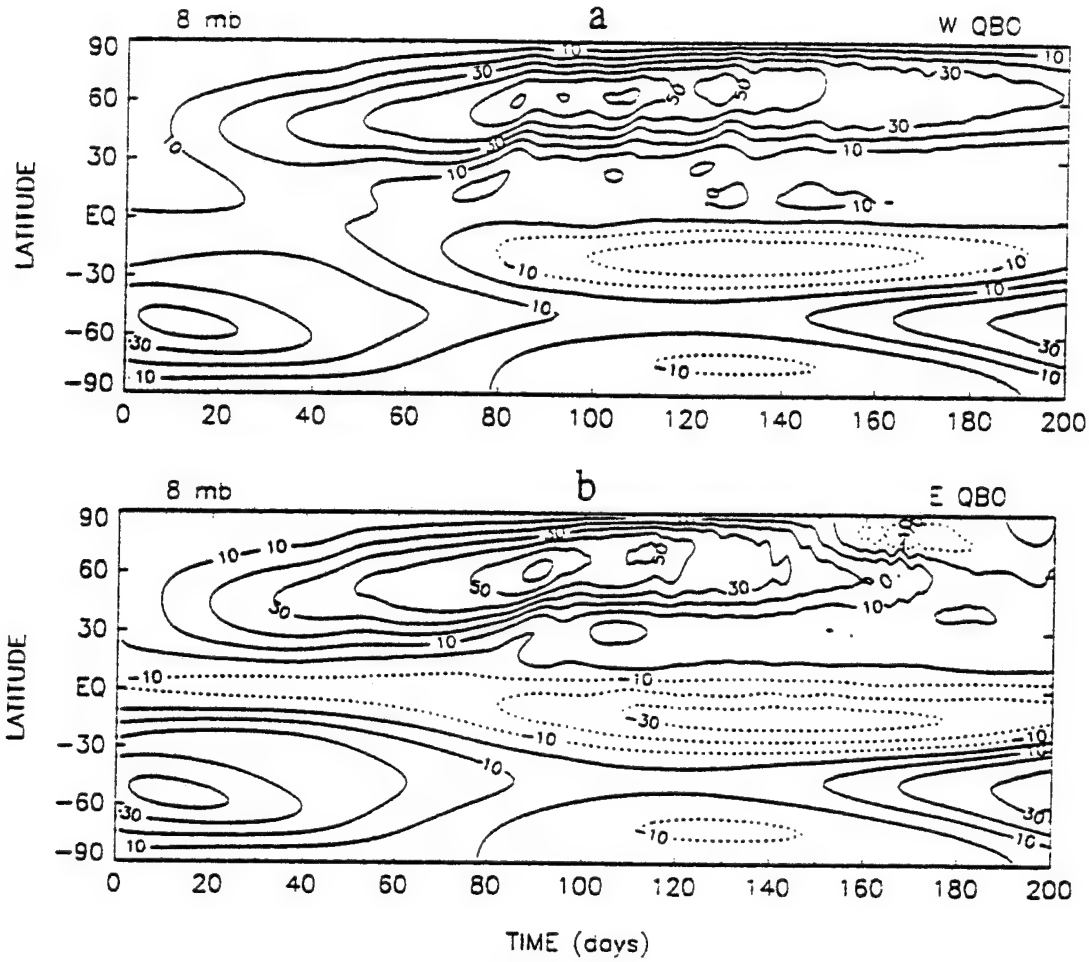


FIG. 4. Latitude–time cross sections of mean zonal wind at 8 mb in seasonal integrations with parameterized wave drag; forcing amplitude ϕ_m = 275 m. (a) West phase; (b) east phase. Model output smoothed with 3-day running mean.

itive (negative) anomaly in polar (middle) latitudes—appeared in most of their experiments with little vertical structure, maximizing in the mesosphere. In Fig. 7, the extratropical anomaly (when significant) maximized in the middle or upper stratosphere and was overlain by a mesospheric anomaly of opposite sign contiguous to the midlatitude half of the stratospheric dipole. The mesospheric response may be a shielding effect, whereby the planetary waves, having caused the stratospheric vortex to weaken or break down, no longer penetrate the mesosphere well enough to overcome the restoring force of radiative relaxation. This interpretation is supported by Figs. 8a,b, showing the latitude–time structure of mean zonal wind at 0.2 mb with ϕ_m = 325 m. The corresponding plots of wavenumber 1 Rossby wave amplitude at 1.0 mb are shown in Figs. 9a,b; Fig. 9c shows the DJF-average difference of wave amplitude. At upper levels it was the *west* phase that was more exposed to Rossby waves, causing a more sustained disruption of the flow than in the east phase. Relaxation of upper-level flow is vaguely analogous to the shielding effect observed in the equatorial QBO (Plumb 1977).

As in the previous subsection, the effect of the equatorial QBO on the extratropical stratosphere could be likened to a shift of the point of maximum sensitivity to ϕ_m, or "fuzzy bifurcation." Figure 10 illustrates the effect schematically for both series, showing the time-averaged mean zonal flow at 70°N, 8 mb, as a function of ϕ_m. The largest W–E difference was obtained in the middle range where the curves are farthest apart.

Due to parameterized wave drag, the mesospheric response was more complicated. Figure 11 shows the time-averaged profiles of mean zonal wind at 70°N from the experiments of Fig. 8. Mean zonal winds were decelerated to a greater degree in the *west* phase, reversing the sign of W–E difference compared to the stratospheric extratropical QBO. Figure 11 also shows analogous profiles obtained from Fig. 2a,b without wave drag. There was no reversal in this case.

Two distinct effects of the wave drag parameterization are therefore apparent when compared to results

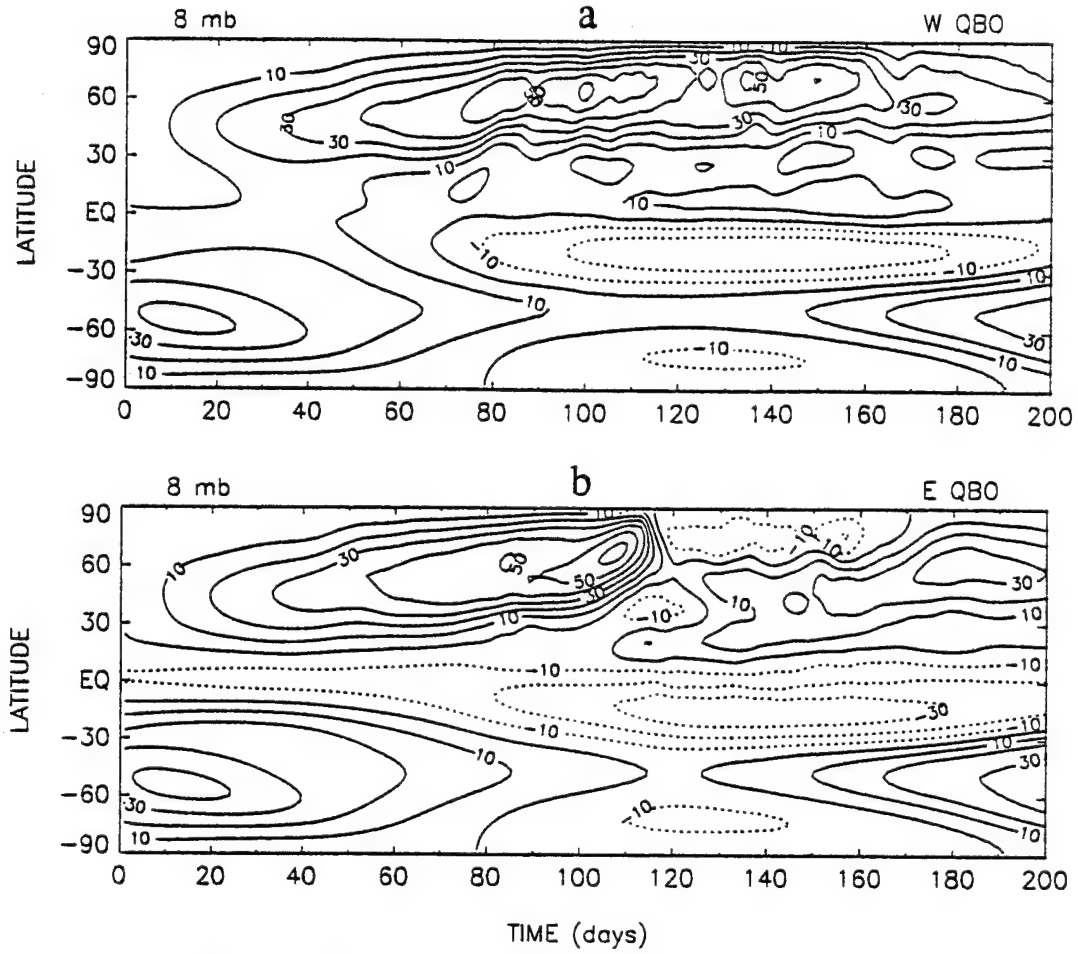


FIG. 5. As in Fig. 4 but with ϕ_m = 300 m. (a) West phase; (b) east phase.

of the previous subsection. First, drag reduced the amplitude of the mesospheric jet to begin with, so any deceleration caused by Rossby waves was necessarily less than that of Fig. 2c. Second, by reducing the amplitude of the jet, the specified drag allowed planetary Rossby waves to propagate to the mesosphere regardless of ϕ_m and enabled smaller values of ϕ_m to cause vortex breakdown. With parameterized wave drag there is no sharp bifurcation between extremes, because the radiatively determined state is no longer admissible as a solution. This is especially relevant in the mesosphere, where planetary Rossby waves cannot penetrate the radiatively determined state (Charney and Drazin 1961; cf. Fig. 2a) but easily propagate into a jet decelerated by unresolved body forces[3] (Fig. 8a) if they are not first absorbed in the stratosphere.

Clear evidence of a reversal of the Holton–Tan oscillation in the mesosphere has not been obtained for lack of observations, although Dunkerton and Baldwin (1991) found a maximum extratropical zonal-wind QBO at 5 mb (with diminished amplitude at 2 and 1 mb). Above 5 mb the polar temperature QBO reversed sign as required by thermal wind balance. Within individual years, observations of mesospheric cooling above stratospheric warmings are fairly common (Quiroz 1969; Labitzke 1972; Hirota and Barnett 1977; Matsuno and Nakamura 1979; Holton 1983; Delisi and Dunkerton 1988).

The sequence of extratropical QBOs in Fig. 7 was not unique to these experiments. A better-resolved series with forcing increments as low as 12.5 m (using a different frictional drag coefficient) also gave a continuous variation of $\Delta\bar{u}$ as a function of ϕ_m. It is noteworthy that the extratropical QBO—when defined over

[3] Gravity wave drag presumably damps planetary Rossby waves in the mesosphere (Miyahara 1985). This effect was ignored, but if included would not alter our conclusion that gravity wave drag, first of all, *enables* planetary wave propagation by preventing a radiatively determined state.

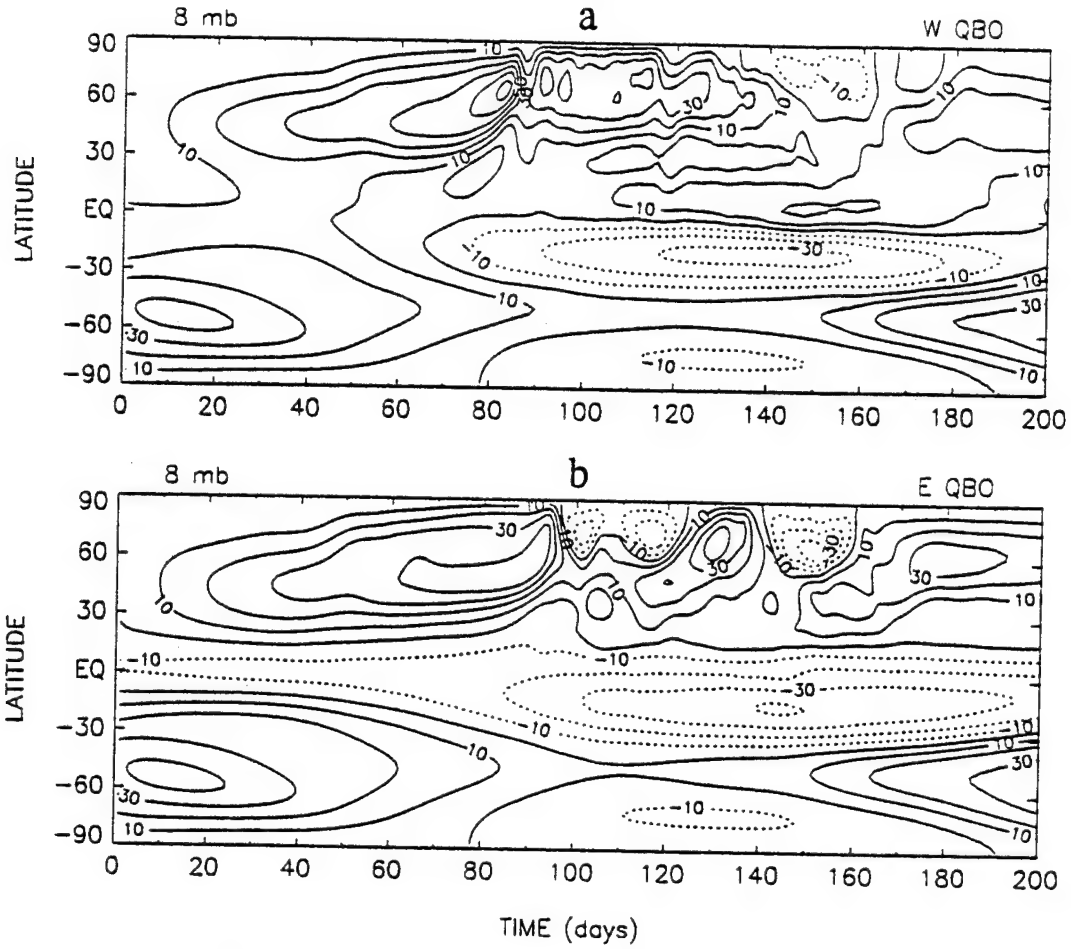


FIG. 6. As in Fig. 4 but with ϕ_m = 350 m. (a) West phase; (b) east phase.

a suitably long time interval, that is, 90 days—depends smoothly on ϕ_m when, in fact, the day-to-day evolution is chaotic after the initial breakdown. This is probably attributable to the averaging interval being long compared to the radiative and mechanical damping times.

Little interannual variability is expected in a strongly damped, nonlinear dynamical system with fixed external forcing (Yoden 1990)—a conjecture that remains to be verified in our model. The point is academic for the terrestrial atmosphere since tropospheric forcings vary from year to year. Composites of the extratropical QBO blend together these interannual variations unrelated to the QBO and are therefore not expected to match precisely the results of any individual experiment as shown in Fig. 7, for example. To account for interannual (non-QBO) variability we constructed a "model composite" of the extratropical QBO averaging together results of the four experiments shown in that figure, which probably represent a typical interannual range of tropospheric forcings. Composites of W–E differences in zonal-mean wind, temperature, and wave amplitude are shown in Figs. 12a–c. The wind and temperature differences agreed reasonably well with observations, both in magnitude and (for temperature) the altitude of sign reversal.

4. Discussion

The sensitivity of extratropical QBO to the amplitude of forced planetary Rossby waves suggests that the Holton–Tan oscillation will vary from year to year, as tropospheric forcing varies. Observations show that the oscillation is not constant, but varies on a decadal timescale (Dunkerton and Baldwin 1992).

Generally speaking, factors that influence the extratropical QBO can be categorized in one of four ways: effects from above (mesosphere), from low latitudes (equatorial stratosphere), in situ anomalies (polar stratosphere), and forcing from below (troposphere).

a. Mesosphere

The mesosphere can affect planetary wave propagation and alter the seasonal development of strato-

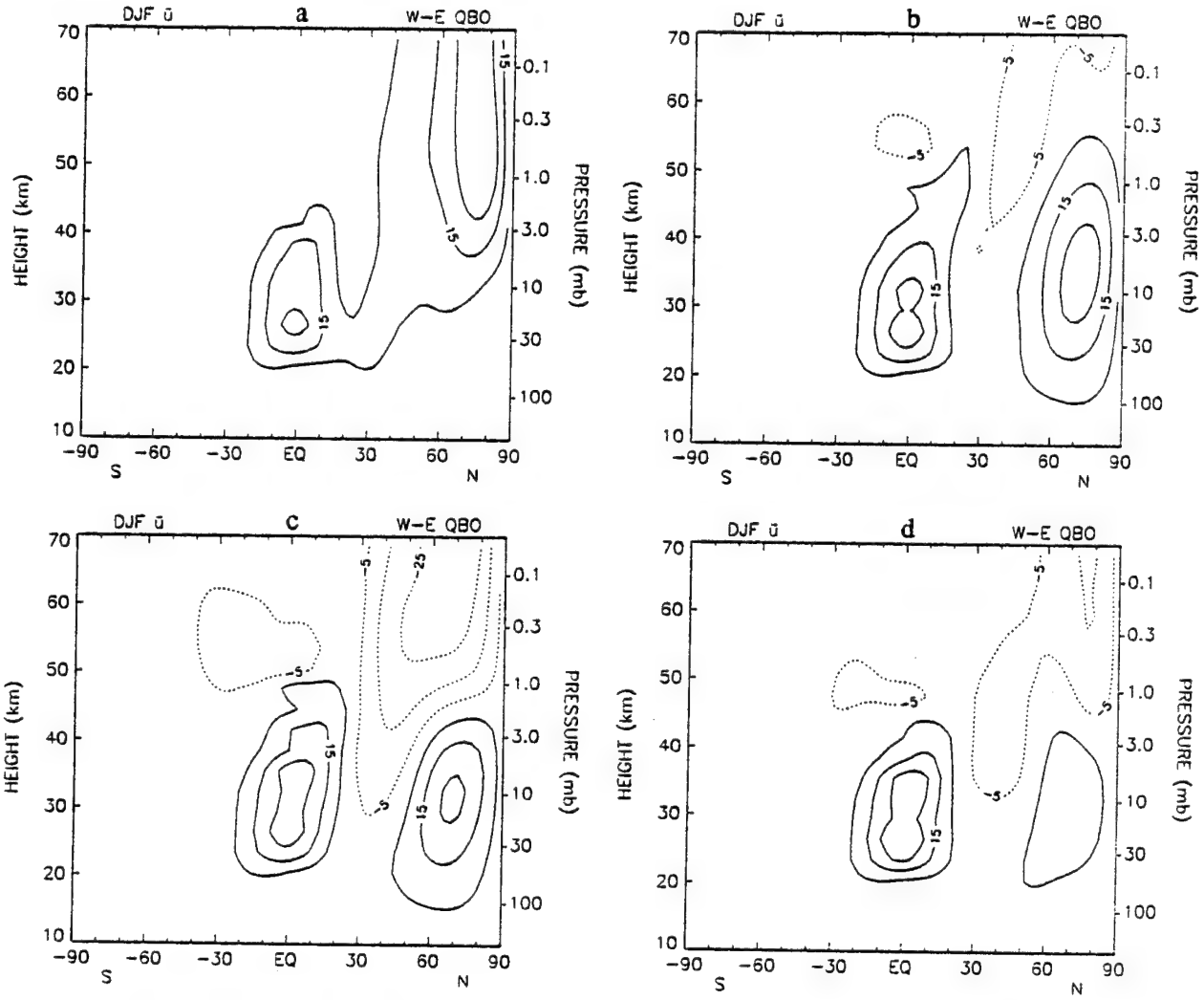


FIG. 7. DJF mean zonal wind difference (west minus east phase) for various forcing amplitudes ϕ_m: (a) 275 m; (b) 300 m; (c) 325 m; (d) 350 m.

spheric circulation. Gravity wave drag illustrates how unresolved body forces in the mesosphere indirectly affect the stratosphere. By retarding the mean flow, gravity waves establish a broad waveguide for planetary Rossby waves that would otherwise be absent. To be sure, planetary waves decelerate the mean flow and thereby maintain their own waveguide. This is true even in experiments without parameterized wave drag. However, the extreme variation between radiatively determined and disturbed states (section 3a) suggests that gravity wave drag, or some other retarding force, is necessary for a realistic seasonal cycle and extratropical QBO.

It is conceivable that other dynamical processes (e.g., barotropic instability) and thermodynamic ones (e.g., anomalous heating due to the solar cycle) could have some downward influence on the stratosphere. Interest in the solar cycle arises from an apparent correlation with the amplitude of extratropical QBO (Labitzke and van Loon 1988; Dunkerton and Baldwin 1992; Kodera 1993). Observations suggest that the December mesospheric jet is anomalously strong near solar maximum (Kodera and Yamazaki 1990). (Whether this association will survive the test of time remains to be seen. Centuries will pass before solar cycle effects can be established with statistical confidence.) Kodera et al. (1991) attempted to force the observed wind variation in a numerical model by altering the solar heating due to ozone. The magnitude of their imposed change was significantly larger than current estimates (by a factor of 10–20). The pattern of December wind anomaly in Kodera and Yamazaki (1990)—fairly deep but restricted to a narrow band

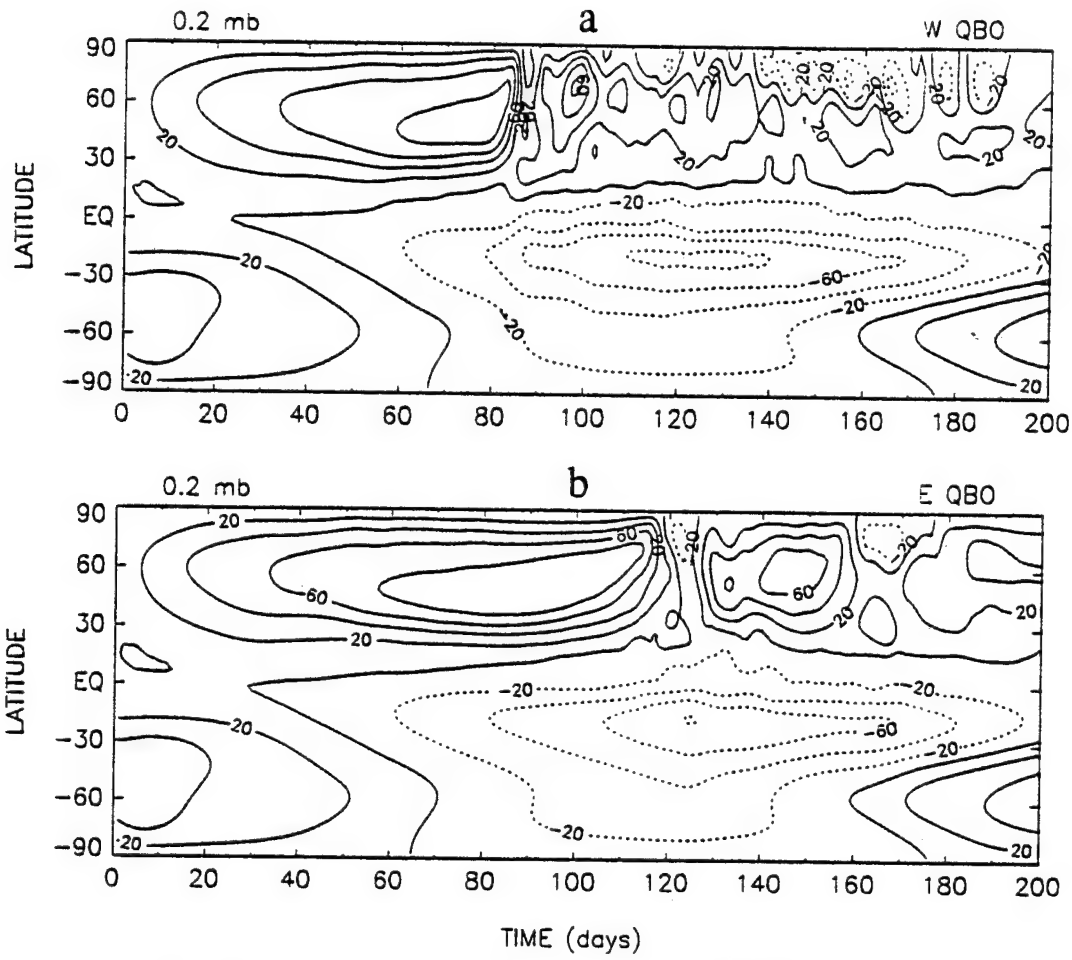


FIG. 8. Latitude–time cross sections of mean zonal wind at 0.2 mb, with forcing amplitude ϕ_m = 325 m. (a) West phase; (b) east phase.

of latitudes—suggests an anomalous body force (not anomalous heating) as the source.[4]

b. *Equatorial stratosphere*

The timing of the QBO will vary from year to year since it is not synchronized with the seasonal cycle. Over several winters the high-latitude flow will "see" different profiles of tropical mean wind. This will modulate the extratropical QBO on a timescale from 5 to 11 years, corresponding to the length of time required for QBO onsets to advance 12 months through the calendar year (Gray and Dunkerton 1990). Further investigation of modulation will require multiyear extensions of the seasonal integrations performed here.

Anomalous heating due to volcanic aerosol in the tropical lower stratosphere was shown by Pitari (1993) to enhance planetary wave propagation in a numerical model.

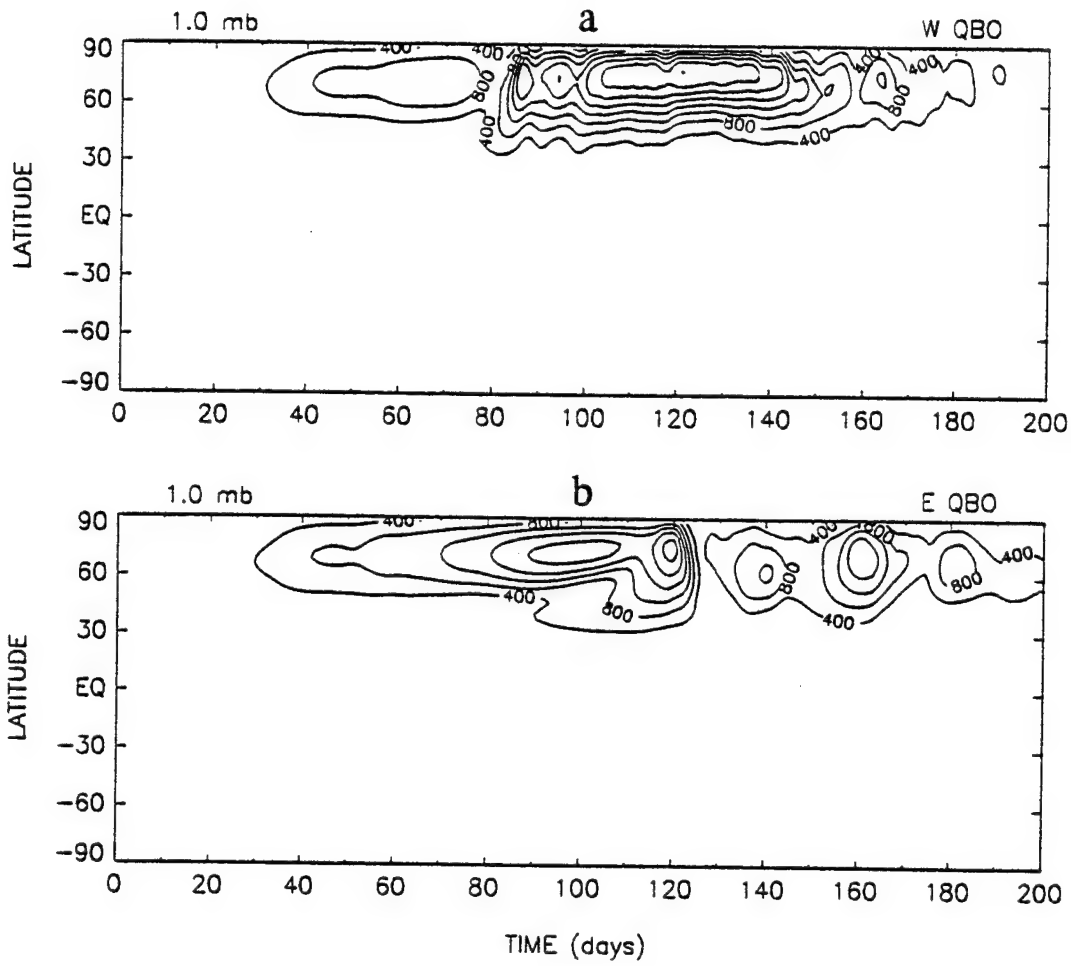
c. *Polar lower stratosphere*

Anomalous radiative effects in the polar lower stratosphere (e.g., due to ozone depletion) can have a positive feedback on planetary wave propagation (Houben 1989, unpublished abstract). Unlike mesospheric anomalies, perturbations in the lower stratosphere alter the mean wind at the point of entry for planetary waves and can therefore have a significant effect even when the imposed changes are small.

d. *Troposphere*

Tropospheric variations affect the stratosphere, not only via forced waves but as the whole "lower bound-

[4] Our results, to be reported elsewhere, demonstrate that it is possible to modify the extratropical QBO if a wind variation like that of Kodera and Yamazaki is imposed on the model. Thermal forcing alone is insufficient since the magnitude of variation is unreasonably large, and thermal forcing is accompanied by nonlocal effects in latitude such as a stronger diabatic circulation and easterlies (hence, relative cooling) near the tropical stratopause (Dunkerton 1991).



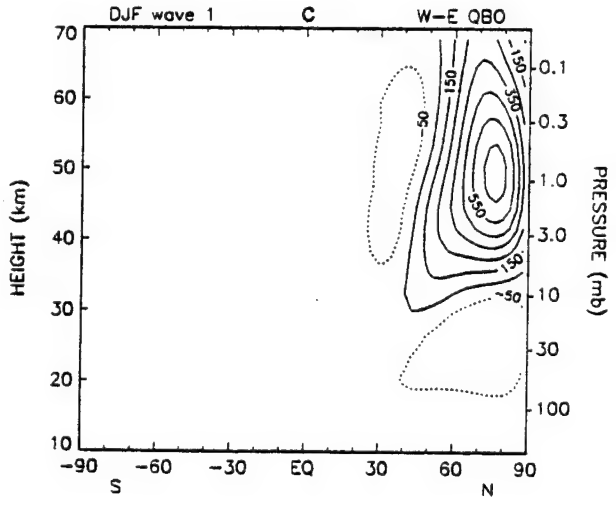


FIG. 9. As in Figs. 8a,b but for geopotential height amplitude of zonal wavenumber 1 at 1.0 mb. (c) DJF W–E difference of wave amplitude.

ary condition'' (including the mean zonal wind) is modified by low-frequency variability. Certain patterns of low-frequency variability in the troposphere are favorably associated with planetary wave excitation and vortex breakdown (Baldwin et al. 1994).

If results of Holton and Dunkerton (1978) are valid in a three-dimensional context, bifurcation of the polar vortex depends also on the specified mean zonal wind at the tropopause. An investigation of the stratospheric response to tropospheric low-frequency

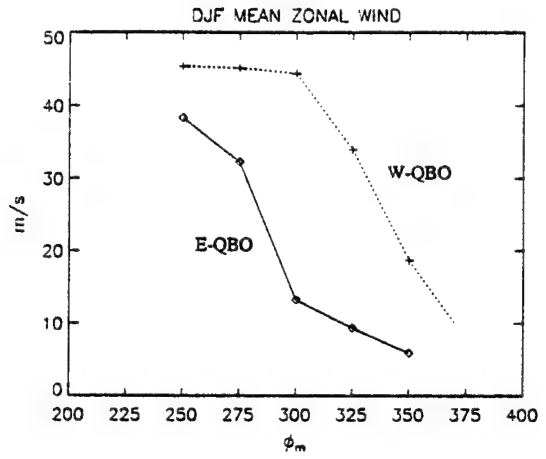


FIG. 10. Values of DJF mean zonal wind at 70°N, 8 mb in several experiments with parameterized wave drag, illustrating the shift of "fuzzy bifurcation" from east to west phase.

variability is in progress, and results will be reported elsewhere.

5. Conclusions

A three-dimensional nonlinear numerical model of the middle atmosphere was used to investigate the sea-

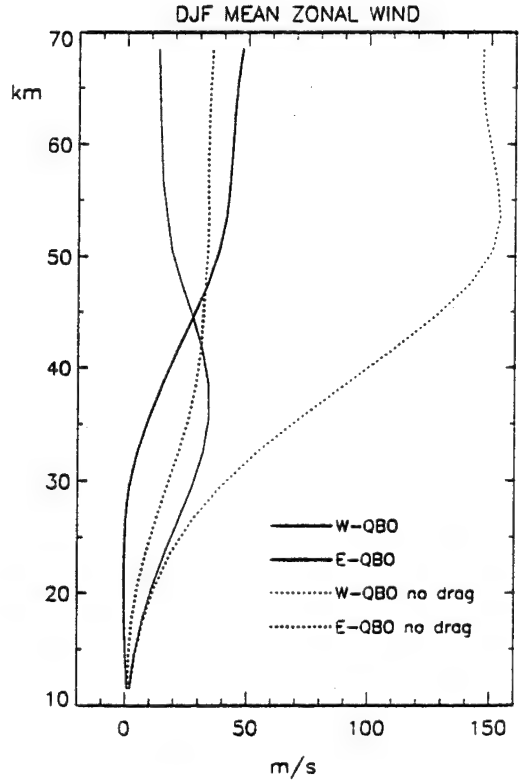


FIG. 11. Vertical profiles of DJF mean zonal wind in experiments with parameterized wave drag (ϕ_m = 325 m), and without parameterized wave drag (ϕ_m = 375 m.)

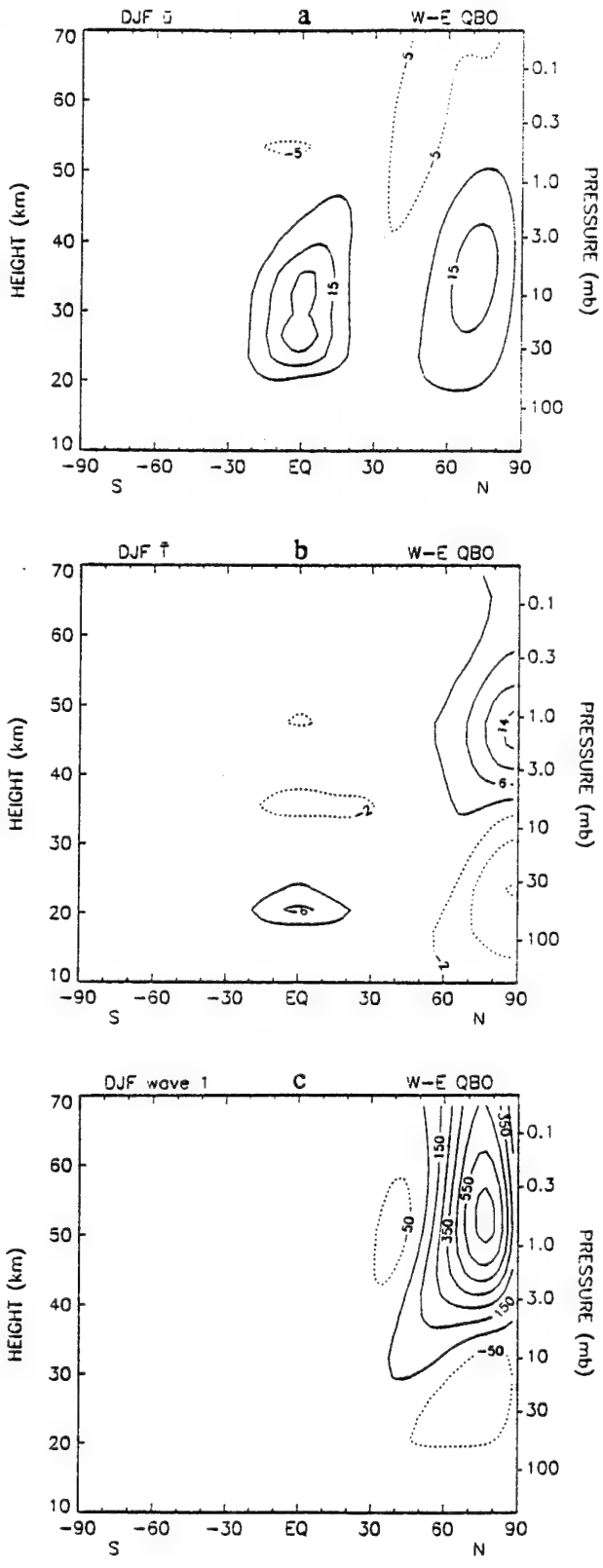


FIG. 12. Model composite of DJF W–E difference combining results from the four experiments of Fig. 7. (a) Mean zonal wind: (b) temperature: (c) geopotential height amplitude of zonal wavenumber 1.

sonal development of the extratropical quasi-biennial oscillation (QBO). Several twin experiments were performed, differing initially only in the phase of equatorial QBO, with identical planetary Rossby waves forced at a tropopause lower boundary. Integrations evolved during winter in a manner roughly consistent with observations, but were sensitive to the lower boundary forcing and the parameterization of radiative cooling and unresolved body force. Stronger radiative forcing (compared to O'Sullivan and Young 1992) heightened the sensitivity of the polar vortex to tropospheric forcing, resulting in a large extratropical QBO when the wave forcing ϕ_m was near a "bifurcation point." As in Holton and Dunkerton (1978), the bifurcation is the value of ϕ_m for which the time-averaged flow is most sensitive to changes in ϕ_m. The effect was most dramatic in perpetual solstice runs where, for certain values of ϕ_m, the equatorial QBO caused the model to cross from a radiatively determined state to a major warming. The sensitivity was reduced but not eliminated in experiments with a seasonal cycle, so it is appropriate to speak of a "fuzzy bifurcation" in this case.

The unrealistically large amplitude of extratropical QBO was reduced substantially by parameterized wave drag, which had two distinct effects. First, maximum wind speeds in the polar vortex were reduced so that any composite difference would be considerably less than the difference between radiatively determined and disturbed states alluded to in the previous paragraph. Second, propagation of Rossby waves was enhanced (Charney and Drazin 1961), making it easier for the polar vortex to break down. Equivalently, less tropospheric forcing was required to achieve a breakdown. This is important in the mesosphere since the radiatively determined state, unhindered by wave drag, precludes vertical propagation of planetary waves.

Realistic extratropical QBOs were obtained for observed values of ϕ_m when excessive westerlies were relaxed to climatological values (Barnett and Corney 1985) using a simple parameterization scheme. Unlike previous model results, but consistent with observations (Dunkerton and Baldwin 1991), the extratropical QBO maximized in the upper stratosphere when realistic ϕ_m were specified. The anomaly reversed sign in the mesosphere, apparently due to a shielding effect as explained in section 3b. Further observations are needed to determine whether the phase reversal is significant.

Numerical experiments demonstrate that the extratropical QBO is sensitive to the amplitude of planetary Rossby waves entering the stratosphere, and requires additional wave drag in the mesosphere to be realistic. Other factors may also affect the oscillation, such as anomalous heating, timing of equatorial QBO, ozone depletion, and tropospheric low-frequency variability. Further investigation of some of these mechanisms may be necessary if we are to understand the observed variation of extratropical QBO.

Finally, our experiments suggest a paradox in tropical–extratropical interaction: how the equatorial QBO, on the one hand, could be shielded from the influence of midlatitude planetary waves while, on the other hand, having a significant influence on these waves in extratropical latitudes. (The simulated west phase of equatorial QBO would be eliminated by planetary wave transport in many cases, apart from a westerly restoring force.) Would a multiyear numerical integration, containing Kelvin waves, Rossby–gravity waves, and midlatitude planetary waves equilibrate to a realistic equatorial QBO, generate an extratropical signal, but remain unaffected at low latitudes by midlatitude waves? We speculate that this may require formation of a region of weak climatological easterlies in the subtropics (Dunkerton and Delisi 1985b). Experiments with more realistic QBO structure and evolution would shed light on the momentum balance of the QBO (as affected by lateral propagation of planetary waves, particularly in the subtropics), the nature of the subtropical "dynamical barrier," and its effect on wave propagation and transport. They would also demonstrate whether the extratropical effects described in this paper can occur under more realistic conditions.

We anticipate that it may soon be feasible to perform a few such integrations on supercomputers.

Acknowledgments. This research was supported by the National Science Foundation Grants ATM-9024392, ATM-9123797, and by the Air Force Office of Scientific Research Contracts F49620-89-C-0051 and F49620-92-C-0033. Numerical computations were performed at the National Center for Atmospheric Research, which is sponsored by the National Science Foundation.

REFERENCES

Baldwin, M. P., and T. J. Dunkerton, 1991: Quasi-biennial oscillation above 10 mb. *Geophys. Res. Lett.*, **18**, 1205–1208.

——, X. Cheng, and T. J. Dunkerton, 1994: Observed correlations between winter-mean tropospheric and stratospheric circulation anomalies. *Geophys. Res. Lett.*, **21**, 1141–1144.

Barnett, J. J., and M. Corney, 1985: Middle atmosphere reference model derived from satellite data. *Middle Atmosphere Program, Vol. 16, Handbook for MAP*. K. Labitzke, J. J. Barnett, and B. Edwards, Eds., 318 pp.

Bojkov, R. D., 1986: The 1979–1985 ozone decline in the Antarctic as reflected in ground based observations. *Geophys. Res. Lett.*, **13**, 1236–1239.

Charney, J. G., and P. G. Drazin, 1961: Propagation of planetary scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83–109.

Dameris, M., and A. Ebel, 1990: The QBO and major stratospheric warmings: A three-dimensional model study. *Ann. Geophys.*, **8**, 79–86.

Delisi, D. P., and T. J. Dunkerton, 1988: Seasonal variation of the semiannual oscillation. *J. Atmos. Sci.*, **45**, 2772–2787.

Dunkerton, T. J., 1991: Nonlinear propagation of zonal winds in an atmosphere with Newtonian cooling and equatorial wavedriving. *J. Atmos. Sci.*, **48**, 236–263.

——, and N. Butchart, 1984: Propagation and selective transmission of internal gravity waves in a sudden warming. *J. Atmos. Sci.*, **41**, 1443–1460.

——, and D. P. Delisi, 1985a: The subtropical mesospheric jet observed by the *Nimbus 7* Limb Infrared Monitor of the Stratosphere. *J. Geophys. Res.*, **90**, 10 681–10 692.

——, and ——, 1985b: Climatology of the equatorial lower stratosphere. *J. Atmos. Sci.*, **42**, 376–396.

——, and M. P. Baldwin, 1991: Quasi-biennial modulation of planetary wave fluxes in the Northern Hemisphere winter. *J. Atmos. Sci.*, **48**, 1043–1061.

——, and M. P. Baldwin, 1992: Modes of interannual variability in the stratosphere. *Geophys. Res. Lett.*, **19**, 49–52.

Fritts, D. C., 1984: Gravity wave saturation in the middle atmosphere: A review of theory and observations. *Rev. Geophys. Space Phys.*, **22**, 275–308.

——, and W. Lu, 1993: Spectral estimates of gravity wave energy and momentum fluxes. Part II: Parameterization of wave forcing and variability. *J. Atmos. Sci.*, **50**, 3695–3713.

Garcia, R. R., and S. Solomon, 1987: A possible relationship between interannual variability in Antarctic ozone and the quasi-biennial oscillation. *Geophys. Res. Lett.*, **14**, 848–851.

Gray, L. J., and T. J. Dunkerton, 1990: The role of the seasonal cycle in the quasi-biennial oscillation of ozone. *J. Atmos. Sci.*, **47**, 2429–2451.

Hartmann, D. L., 1983: Barotropic instability of the polar night jet stream. *J. Atmos. Sci.*, **40**, 817–835.

Hirota, I., and J. J. Barnett, 1977: Planetary waves in the winter mesosphere—preliminary analysis of the Nimbus 6 PMR results. *Quart. J. Roy. Meteor. Soc.*, **103**, 487–498.

Holton, J. R., 1983: The influence of gravity wavebreaking on the general circulation of the middle atmosphere. *J. Atmos. Sci.*, **40**, 2497–2507.

——, and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.*, **33**, 2218–2225.

——, and T. J. Dunkerton, 1978: On the role of wave transcience and dissipation in stratospheric mean flow vacillations. *J. Atmos. Sci.*, **35**, 740–744.

——, and H.-C. Tan, 1980: The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb. *J. Atmos. Sci.*, **37**, 2200–2208.

——, and W. M. Wehrbein, 1981: A further study of the annual cycle of the zonal mean circulation in the middle atmosphere. *J. Atmos. Sci.*, **38**, 1504–1509.

——, and ——, 1982: The quasi-biennial oscillation in the Northern Hemisphere lower stratosphere. *J. Meteor. Soc. Japan*, **60**, 140–148.

——, and J. Austin, 1991: The influence of the equatorial QBO on sudden stratospheric warmings. *J. Atmos. Sci.*, **48**, 607–618.

Kodera, K., 1993: Quasi-decadal modulation of the influence of the equatorial quasi-biennial oscillation on the north polar stratospheric temperatures. *J. Geophys. Res.*, **98**, 7245–7250.

——, and K. Yamazaki, 1990: Long-term variation of upper stratospheric circulation in the Northern Hemisphere in December. *J. Meteor. Soc. Japan*, **68**, 101–105.

——, M. Chiba, and K. Shibata, 1991: A general circulation model study of the solar and QBO modulation of the stratospheric circulation during the Northern Hemisphere winter. *Geophys. Res. Lett.*, **18**, 1209–1212.

Labitzke, K., 1972: Temperature changes in the mesosphere and stratosphere connected with circulation changes in winter. *J. Atmos. Sci.*, **29**, 756–766.

——, 1982: On the interannual variability of the middle stratosphere during the northern winters. *J. Meteor. Soc. Japan*, **60**, 124–139.

——, 1987: Sunspots, the QBO, and the stratospheric temperature in the north polar region. *Geophys. Res. Lett.*, **14**, 535–537.

——, and H. van Loon, 1988: Associations between the 11-year solar cycle, the QBO, and the atmosphere. Part I: The troposphere and stratosphere in the Northern Hemisphere winter. *J. Atmos. Terr. Phys.*, **50**, 197–206.

Lait, L. R., M. R. Schoeberl, and P. A. Newman, 1989: Quasi-biennial modulation of the Antarctic ozone depletion. *J. Geophys. Res.*, **94**, 11 559–11 571.

Lindzen, R. S., 1981: Turbulence and stress due to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86C**, 9707–9714.

Matsuno, T., and K. Nakamura, 1979: The Eulerian and Lagrangian mean meridional circulations in the stratosphere at the time of a sudden warming. *J. Atmos. Sci.*, **36**, 640–654.

Miyahara, S., 1985: Suppression of stationary planetary waves by internal gravity waves in the mesosphere. *J. Atmos. Sci.*, **42**, 100–107.

O'Sullivan, D., and M. L. Salby, 1990: Coupling of the quasi-biennial oscillation and the extratropical circulation in the stratosphere through planetary wave transport. *J. Atmos. Sci.*, **47**, 650–673.

——, and R. E. Young, 1992: Modeling the quasi-biennial oscillation's effect on the winter stratospheric circulation. *J. Atmos. Sci.*, **49**, 2437–2448.

Pitari, G., 1993: A numerical study of the possible perturbation of stratospheric dynamics due to Pinatubo aerosols: Implications for tracer transport. *J. Atmos. Sci.*, **50**, 2443–2461.

Plumb, R. A., 1977: The interaction of two internal waves with the mean flow: Implications for the theory of the quasi-biennial oscillation. *J. Atmos. Sci.*, **34**, 1847–1858.

Quiroz, R. S., 1969: The warming of the upper stratosphere in February 1966 and the associated structure of the mesosphere. *Mon. Wea. Rev.*, **97**, 541–552.

Randel, W. J., 1992: Global Atmospheric Circulation Statistics, 1000-1 mb. NCAR Tech. Note NCAR/TN-366 + STR, Boulder, CO.

Salby, M. L., D. O'Sullivan, R. R. Garcia, and P. Callaghan, 1990: Air motions accompanying the development of a planetary wave critical layer. *J. Atmos. Sci.*, **47**, 1179–1204.

Shine, K. P., 1987: The middle atmosphere in the absence of dynamical heat fluxes. *Quart. J. Roy. Meteor. Soc.*, **113**, 603–633.

van Loon, H., and K. Labitzke, 1987: The southern oscillation. Part V: The anomalies in the lower stratosphere of the Northern Hemisphere in winter and a comparison with the quasi-biennial oscillation. *Mon. Wea. Rev.*, **115**, 357–369.

Wallace, J. M., and F.-C. Chang, 1982: Interannual variability of the wintertime polar vortex in the Northern Hemisphere middle stratosphere. *J. Meteor. Soc. Japan*, **60**, 149–155.

Yoden, S., 1987a: Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.*, **44**, 1723–1733.

——, 1987b: Dynamical aspects of stratospheric vacillations in a highly truncated model. *J. Atmos. Sci.*, **44**, 3683–3695.

——, 1987c: A new class of stratospheric vacillations in a highly truncated model due to wave interference. *J. Atmos. Sci.*, **44**, 3696–3709.

——, 1990: An illustrative model of seasonal and interannual variations of the stratospheric circulation. *J. Atmos. Sci.*, **47**, 1845–1853.

——, and K. Ishioka, 1993: A numerical experiment on the breakdown of a polar vortex due to forced Rossby waves. *J. Meteor. Soc. Japan*, **71**, 59–72.

CISK and Evaporation-Wind Feedback with Conditional Heating on an Equatorial Beta-Plane

Francis X. Crum and Timothy J. Dunkerton

CISK and Evaporation-Wind Feedback with Conditional Heating on an Equatorial Beta-Plane

By Francis X. Crum and Timothy J. Dunkerton

Northwest Research Associates, P.O. Box 3027, Bellevue, WA 98009, U.S.A.
(Manuscript received 5 January 1993, in revised form 12 October 1993)

Abstract

The effect of conditional heating on unstable Kelvin waves generated by CISK and evaporation-wind feedback was investigated with multi-level models on the equatorial beta-plane. In agreement with previous analytic results, the scale-selection catastrophe of linear heating is modified but not eliminated by conditional heating. There is only one region of upward motion, but it contracts to the smallest zonal scale allowed by diffusion or model resolution. The pathology is most severe in the CISK-unstable case. With evaporation-wind feedback, Kelvin-wave instability is possible apart from CISK; growth rate in this case increases slowly with decreasing diffusion, but contraction is still observed. We note incidentally that eastward moving evaporation-wind modes can be generated on the beta-plane with no mean wind, contrary to previous assumptions requiring an easterly background flow.

While evaporation-wind feedback is an improvement over CISK, as far as scale selection is concerned, zonal phase speeds remain too large compared to the observed propagation of tropical intraseasonal oscillations and superclusters, unless (1) the parameter q controlling the strength of moisture convergence feedback lies close to the CISK-neutral point, (2) the parameter A controlling the strength of evaporation-wind feedback is small, and (3) the heating profile is assumed proportional to vertical velocity. Small positive phase speeds are obtained under these conditions, for example, in a two-level model.

1. Introduction

Theoretical models of the tropical intraseasonal oscillation (TIO: Madden and Julian, 1971, 1972; Knutson and Weickmann, 1987; Rui and Wang, 1990) have often relied on CISK (Lau and Peng, 1987; Chang and Lim, 1988) or evaporation-wind feedback (Emanuel, 1987; Neelin *et al.*, 1987) to generate unstable waves. Similar mechanisms were shown to be important in GCMs containing intraseasonal oscillations (Hayashi and Sumi, 1986; Swinbank *et al.*, 1988; Lau *et al.*, 1988). Nevertheless, these 'explanations' of TIO are incomplete, as they fail to reproduce the slow eastward phase speed, horizontal structure, and observed hierarchy of convection within intraseasonal oscillations. Furthermore, the exponentially-growing modes are pathological, preferring the smallest zonal scale. This is true whether linear (positive and negative) or conditional (positive-only) heating is assumed.

We recently documented the scale-selection catastrophe analytically with conditional heating (Crum and Dunkerton, 1992; hereafter CD) and now extend the demonstration to multi-level models on an equatorial beta-plane (Section 2b). Our conclusion concerning CISK modes is in agreement with a statement by Yoshizaki (1991) who, apart from this, did not demonstrate the catastrophe nor investigate the role of evaporation-wind feedback. Because linear theory predicts a weaker catastrophe with evaporation-wind feedback, we were motivated to explore the effect of conditional heating and beta-plane geometry (Section 2c). In this context we demonstrate that it is possible to avoid the linearization of evaporation and still obtain an eastward-propagating instability. Variation of zonal phase speed as a function of moisture convergence and evaporation-wind feedback is briefly discussed in Section 2d.

©1994, Meteorological Society of Japan

2. Model results

a. *Numerical model*

We consider linear, hydrostatic perturbations from a time-independent basic state on an equatorial beta-plane. A three-dimensional primitive-equation model was formulated in pressure coordinates with two or five vertical levels (Charney-Phillips grid), finite differences in pressure and latitude, and a Fourier spectral representation in longitude. Solu-

tions were periodic in x. dependent variables were set to zero at the sidewalls ($\pm 60°$ latitude), and vertical boundary conditions were $\omega = 0$ at $p = 0$ and $p = p_{00} = 1000$ mb. Static stability was derived from Houghton (1986) at 10°N. A small second-order horizontal diffusion was inserted on the rhs of horizontal momentum and temperature equations (which nevertheless proved to be essential for well-resolved simulations). Precipitation was parameterized in terms of total moisture entering a column, including that from surface evaporation:

$$P = \begin{cases} -\int_0^{p_{00}} \nabla \cdot (\boldsymbol{V} q)\, dp/g + E_{00}, & \text{if} \geq 0: \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

where P is precipitation in units of $\mathrm{kg\,m^{-2}\,s^{-1}}$, q is water vapor mixing ratio, $\boldsymbol{V}$ is horizontal velocity, and E_{00} is surface evaporation is units of P. Spatially constant q was assumed, non-zero only in the lowest model layer, and heating was distributed vertically by a specified profile (CD). In the five-level model, our heating profile η was defined as (η_{200}, η_{400}, η_{600}, η_{800}) $\propto$ (1/2, 1/2, 2, 1) where subscripts refer to pressure level. This profile generates the 'slow CISK' mode of Chang and Lim (1988).

Readers familiar with CISK will recall that the specified heating profile is the mathematical device responsible for vertical mode-coupling; in unstable CISK, two stable modes have coalesced to form an unstable mode (CD). It should be noted that by specifying the heating profile *a priori*, instead of making it proportional to vertical velocity, zonal propagation of CISK modes can occur (as explained in Appendix A of CD). The second alternative implies nonpropagating convection unless additional sources of heat are present, such as might be due to evaporation-wind feedback.

We limit our application of the five-level model primarily to CISK modes obtained with a specified heating profile. To better approximate the concept of reduced moist static stability in evaporation-wind feedback (Yano and Emanuel, 1991) a two-level model was used following Neelin *et al.* (1987). In this case there is no heating 'profile' as such; instead, $\eta \propto 1$ at a single internal half-level, tantamount to a reduced moist static stability (or convective instability in unstable CISK). From this discussion it can be seen that there is an important conceptual and practical distinction to be made between (1) the use of a specified heating profile proportional to vertical velocity at cloud base and (2) the assumption of heating proportional to vertical velocity at all levels. We return to this in Section 2d when discussing the zonal phase speed of unstable modes.

In what follows, 'conditional' heating refers to the nonlinearity introduced by Eq. (2.1). 'Linear' heating employs the upper formula of Eq. (2.1) irrespective of its sign. In both cases, the problem is dynamically linear, independent of solution amplitude, and exponentially-growing normal modes are possible.

b. Wave-CISK

In this case, evaporation is ignored in Eq. (2.1). With our specified heating profile, the third and fourth internal modes coalesce near $q = 4\,\mathrm{g\,kg^{-1}}$ to form an unstable propagating mode with phase speed 15–20 $\mathrm{ms^{-1}}$. We found similar behavior in models with many more vertical levels. Due to coalescence, vertical structure approximates a 'modified first-internal mode' with no internal zero-crossings of vertical velocity ω. This property is desirable observationally, although the 'slow-CISK' heating profile generates outflow too low (near 400 mb), while phase speed remains too large. Taller heating profiles cause faster propagation. Our immediate intent is not to dwell on these defects but to demonstrate the importance of spatial filtering in zonal scale selection. With linear heating, analytic Kelvin wave solutions are nondispersive so that growth rate is proportional to zonal wavenumber times the imaginary part of the eigenvalue (complex phase speed)—as is generally true in the long-wave approximation. With conditional heating, the scale-selection catastrophe is modified but not eliminated. Only one region of upwelling is preferred, but fastest growth occurs at the smallest zonal scale. CD demonstrated this pathology analytically using simple two- and three-level models in the equatorial x–z plane.

The same is evidently true on an equatorial beta-plane, when the number of vertical levels is increased. Figure 1 shows the growth rate of stable CISK modes with conditional heating, $q = 8\,\mathrm{g\,kg^{-1}}$, obtained numerically with the five-level beta-plane model described in Section 2a. The initial condition was a Gaussian temperature profile in latitude and longitude, maximum in mid-troposphere. Various curves refer to different model resolutions (M, N) where M is the number of zonal harmonics and N is the number of latitudinal grid points in one hemisphere. The dotted line shows the width of the region of upward motion for $(M, N) = (128, 44)$. Results of Fig. 1 are exactly reminiscent of CD's Fig. 6, showing that only a few of the simulations are 'well-resolved' (independent of resolution). Coarse resolution severely underestimates growth rate at low diffusion. In well-resolved simulations, wet-region width decreases catastrophically with diffusion as in the analytic solutions of CD.

Horizontal structure of these solutions is essentially a Kelvin wave with negligible meridional velocity. The vertical structure is a modified first internal mode similar to Lau and Peng (1987).

As an aside, we note that gravity waves do not dominate the simulated modes. In linear CISK, gravity waves grow slightly faster than Kelvin waves

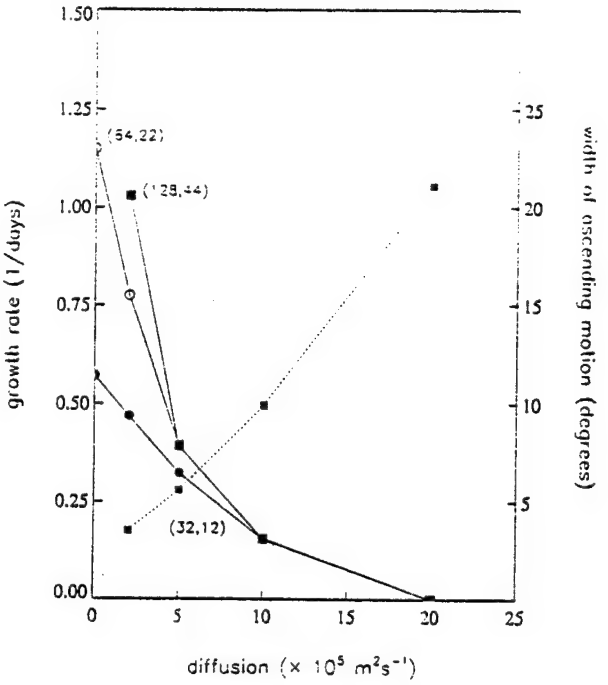


Fig. 1. CISK growth rates in the five-level model as a function of diffusion for $q = 8$ g kg^{-1}. Solid circles, open circles, and squares denote simulation with $(M, N) = (32, 12)$, (64,22), and (128,44) respectively. The e-folding width of ascending motion is shown by the dashed line for $(M, N) = (128, 44)$.

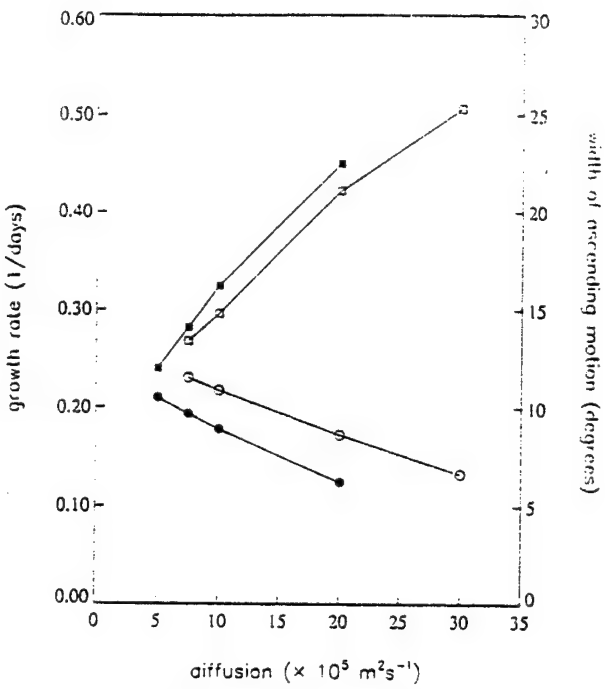


Fig. 2. Evaporation-wind feedback growth rates in the two-level model as a function of diffusion for $q = 14$ g kg^{-1} (CISK-stable case) and $A = 10.0 \times 10^{-6}$ kg m^{-2} s^{-1} (m/s)$^{-1}$. Solid and open circles denote simulations with no mean wind and easterly mean wind, respectively. Open and solid squares show the e-folding width of ascending motion. Results are shown for $(M, N) = (128, 44)$.

(Crum and Stevens, 1983). A series of tests without the y-component of diffusion demonstrated that it was the conditional heating and x-component of diffusion, and not the damping of oscillatory structure in latitude, that suppressed gravity waves. The viscous beta-plane result is interesting because, even with small diffusion, conditional heating leads to a narrow, propagating convergence zone perpendicular to the equator rather than a circularly-symmetric singularity as assumed in the f-plane solutions of Charney and Eliassen (1964).

c. Evaporation-wind feedback

In this case evaporation in Eq. (2.1) is parameterized according to the standard bulk formula

$$E_{00} = \rho_s c_D (q^* - q) |\boldsymbol{V}_s| \tag{2.2}$$

where ρ_s is surface air density, c_D a dimensionless coefficient, q^* is saturation mixing ratio at the surface, and $\boldsymbol{V}_s$ is horizontal surface velocity. Following Neelin *et al.* (1987), if velocity is linearized about a mean zonal flow, the evaporation perturbation may be written

$$E_{00} = A u_s' \mathrm{sgn}(\bar{u}_s) \tag{2.3}$$

where $A = \rho_s c_D (q^* - q)$ is assumed constant, u_s is the zonal velocity component, and the overbar and prime denote zonal-mean and perturbation, respectively. According to the linearized expression Eq. (2.3), an easterly (westerly) perturbation superposed on an easterly (westerly) mean flow acts to enhance heating and vice versa. In the numerical model, horizontal surface velocities were approximated using velocity at the lowest level.

It was shown by CD that CISK-neutral instabilities with evaporation-wind feedback and conditional heating exhibit a weaker scale-selection catastrophe that CISK-unstable modes (weaker in the sense that although wet-region width contracts to zero, growth rate asymptotes to a constant as $\nu \to 0$). Similar results were obtained with the two-level beta-plane model as shown in Fig. 2. For these simulations we chose $q = 14$ g kg^{-1} (below the CISK-neutral point without evaporation-wind feedback) and $A = 10.0 \times 10^{-6}$ kg m^{-2} s^{-1} (m/s)$^{-1}$. This value of A corresponds roughly to an evaporation rate $\sim$4 mm/day induced by a surface perturbation of 5 ms^{-1}. According to Fig. 2, growth rate increases slowly as diffusion decreases, and the width of wet region decreases, as in CD. (The e-folding width of vertical velocity was plotted in Figs. 1, 2 instead of the width

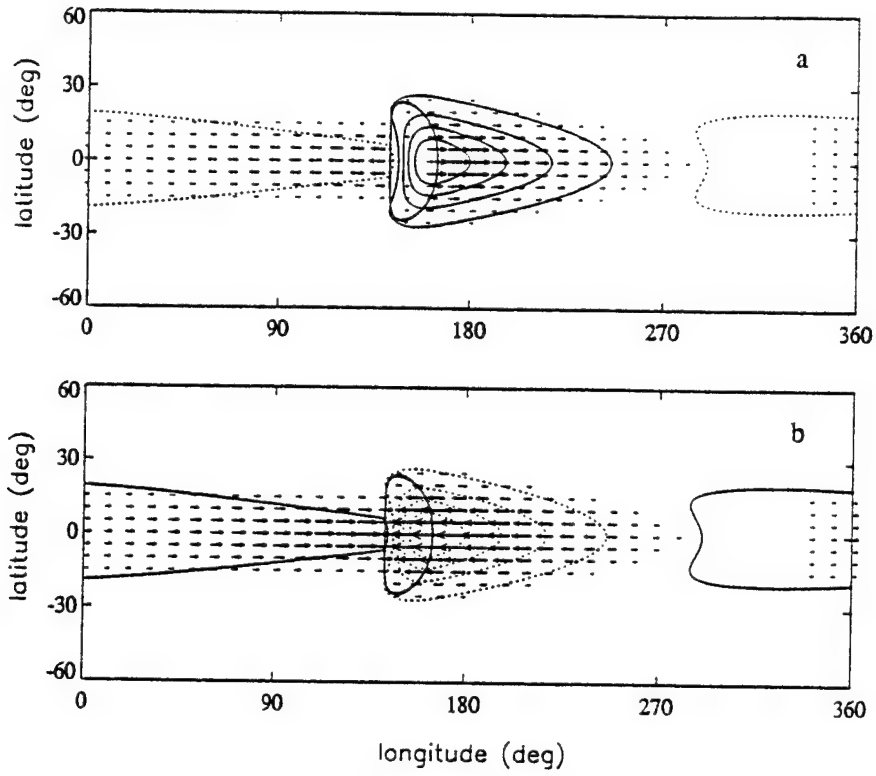


Fig. 3. Structure of unstable evaporation-wind mode with conditional heating and easterly mean flow in the two-level model using $q = 14\,\mathrm{g\,kg^{-1}}$, $A = 10 \times 10^{-6}\,\mathrm{kg\,m^{-3}}$, and $(M, N) = (128, 44)$. Geopotential (thin solid and dashed lines) and horizontal velocity vectors are shown at 250 mb (a) and 750 mb (b). The region of positive heating greater than 10 % of its maximum value lies inside the thick solid line. The horizontal diffusion in this case was $1 \times 10^6\,\mathrm{m^2\,s^{-1}}$.

of positive ω because, in the case of evaporation-wind feedback, there was a broad, zonally asymmetric 'tail' of small ω extending far to the west of maximum convergence.) We would like to explore the small diffusivity limit more rigorously. However, analytic progress is very difficult even in two dimensions (CD) and well behaved numerical solutions require some finite (and not necessarily small) amount of diffusivity.

As shown in Fig. 2, the linearized form Eq. (2.3) is unnecessary for evaporation-wind feedback. Nearly identical eastward-propagating modes are obtained using the exact formula Eq. (2.2) with mean flow neglected. (See also Xie, *et al.*, 1993). The following discussion may help explain this result. A key feature of the mechanism—using Eq. (2.3) with $\bar{u} < 0$—is that warming occurs in the region of perturbation surface easterlies, east of the temperature maximum. This causes eastward propagation. It is irrelevant that perturbation westerlies cause cooling to the west of the convergence maximum when linear heating is assumed. If we now return to the exact formula Eq. (2.2) and set $\bar{u} = 0$, the situation just described would instead cause warming in regions of surface easterlies and westerlies. The direction of propagation may appear ambiguous, were it not for an asymmetric response on the equatorial beta-plane (Gill, 1980). The response is to heating that is propagating and due primarily to a superposition of convergences from propagating Kelvin and Rossby waves. The convergence due to the Kelvin waves is stronger and, consequently, the strongest heating moves eastward. If we now think in terms of the circulation induced by this eastward moving heating, the winds associated with the Kelvin response to the east are stronger than the winds associated with the Rossby response to the west. The WISHE mechanism therefore favors an eastward moving unstable mode. There are two crucial points: the relative strength of the eastward moving heating and the selective enhancement of such heating by WISHE. Thus, eastward propagation is possible on the beta-plane when $\bar{u} \equiv 0$. In fact, solutions are nearly identical to those obtained with Eq. (2.3) since the Kelvin response dominates that case as well.

Horizontal structure, shown in Fig. 3, is essentially an eastward-propagating Kelvin wave as in unstable CISK. A slight difference is that heating extends eastward into the region of surface easterlies due to the contribution from Au'_s. The solution with no mean wind is similar, except for a westward extension of heating into the region of surface westerlies to the west (Fig. 4). Using a westerly mean flow with conditional heating produces a hy-

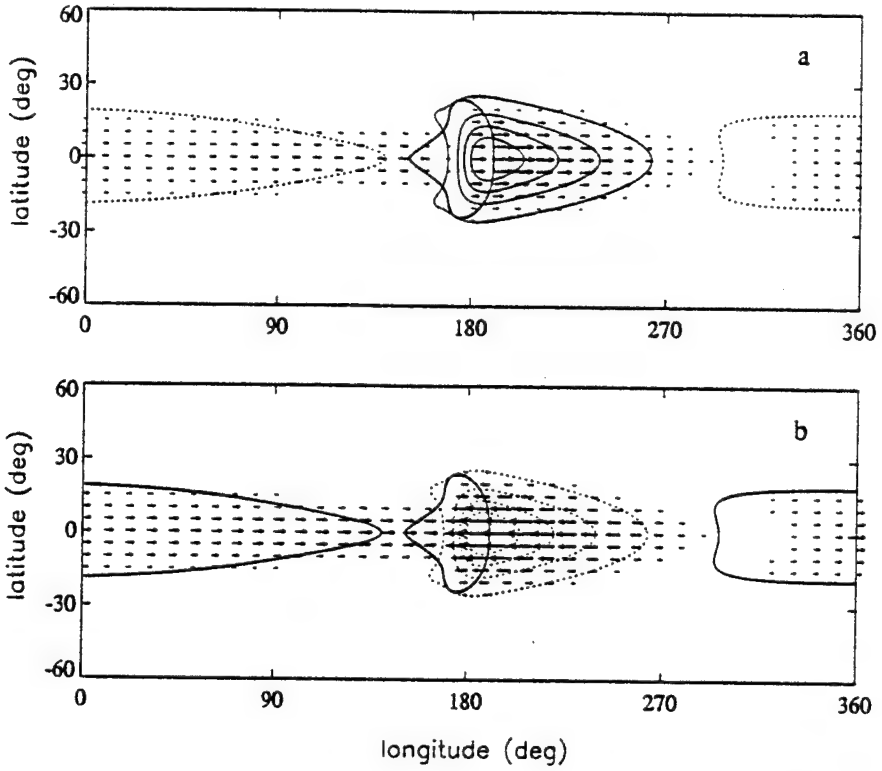


Fig. 4. Horizontal structure as in Fig. 3, but with no mean flow.

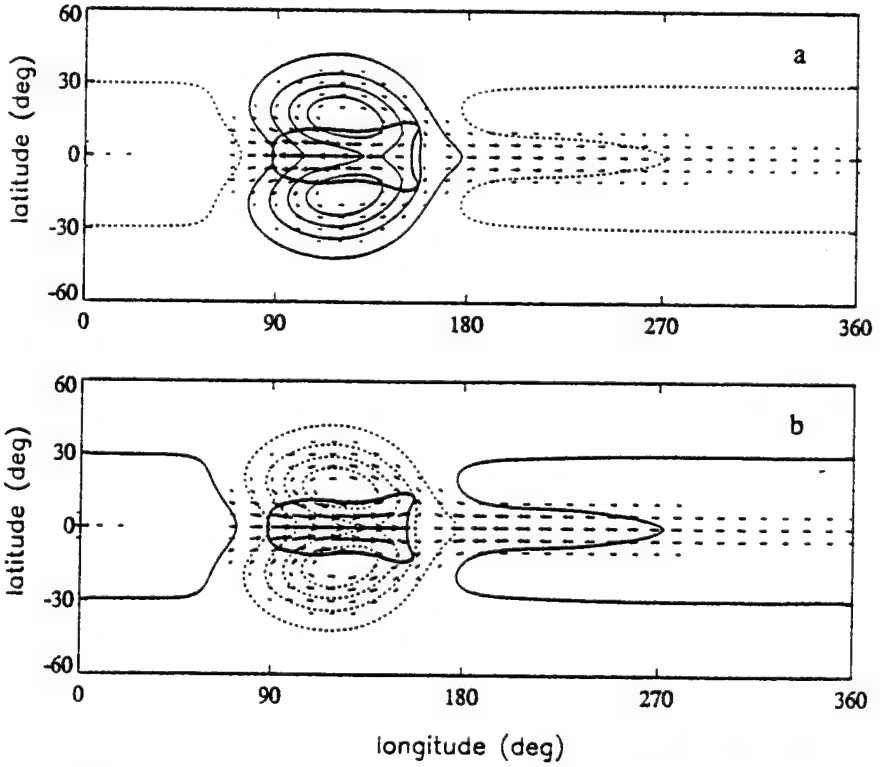


Fig. 5. Horizontal structure as in Fig. 3, but with westerly mean flow.

brid Kelvin/Rossby structure propagating westward (Fig. 5). The hybrid structure owes its existence to the fact that perturbation surface westerlies enhance evaporation to the west of the heating, dragging the entire disturbance westward. Not surprisingly, the Rossby component of the response is considerably larger than in the previous two examples.

Solutions were also obtained with evaporation-

Table 1. Intrinsic phase speeds, $\hat{c}$ (ms^{-1}), and e-folding times, τ (days), of evaporation-wind feedback modes with an easterly basic state and conditional heating as a function of A ($kgm^{-2}\,s^{-1}$ $(m/s)^{-1}$) and q (g/kg). Actual phase speeds relative to the ground can be obtained by adding an easterly flow to the values in the table. Values are displayed as an ordered pair ($\hat{c}$, τ). Results were obtained from the numerical model using $(M, N) = (128, 44)$ and $\nu = 2.5 \times 10^5\ m^2\,s^{-1}$. The CISK neutral value in the absence of evaporation-wind feedback is $q = 19.1$ g/kg. A dash (—) entry means that no unstable mode could be identified for that set of parameters.

	$q = 8$	$q = 14$	$q = 19.1$	$q = 19.5$
$A = 15 \times 10^5$	45.3, 2.6	31.8, 2.4	12.8, 1.2	5.8, 0.6
$A = 10 \times 10^5$	45.2, 4.1	31.6, 3.9	12.7, 2.3	4.5, 0.7
$A = 5 \times 10^5$	45.2, 10.7	31.5, 19.2	—	2.6, 0.8

wind feedback in the five-level model, with heating proportional to the sum of low-level evaporation and moisture convergence and using a specified vertical profile. Results were similar to those of the two-level model (not shown), with certain differences as elucidated in the next subsection. Unlike unstable CISK modes, evaporation-wind feedback in a multi-level model destabilizes individual vertical modes without coalescence. Consequently, zonal phase speeds are similar to those of free modes, and are generally larger than the propagation speed of CISK instabilities (which are already greater than observed TIO speeds). Examples of this behavior were shown in Appendix B of CD. However, we doubt that this generalization of evaporation-wind feedback to a multi-level model—with specified vertical heating profile as in wave-CISK—adequately represents the moist entropy budget as described by Yano and Emanuel (1991). In any case, the phase speeds of unstable Kelvin waves are undesirably large, so we do not discuss these modes further.

d. Effective moist static stability

As noted in the Introduction, theories of the TIO that rely on CISK or evaporation-wind feedback fail to reproduce the slow eastward propagation, horizontal structure (including subtropical gyres), and hierarchy of convection within intraseasonal oscillations. The last two observations suggest the importance of latitudinal shear (Dunkerton, 1990) and group velocity of synoptic-scale waves (Liebmann and Hendon, 1990), and show that a purely Kelvin-wave interpretation of the TIO is inadequate.

Concerning the first point, a general shortcoming of CISK-unstable Kelvin waves in a multi-level model with specified vertical heating profile is that moist stable modes (contiguous to dry free modes) coalesce to form an instability at unrealistically high phase speeds similar to those of dry modes. When evaporation-wind feedback is added, CISK-stable instabilities remain fast (CD).

An exception to this statement occurs in the two-level model destabilized by evaporation-wind feedback. In this case phase speeds can be made arbitrarily small, according to linear theory, by approaching the CISK-neutral point from below and simultaneously reducing A. Reduction of phase speed (and of latitudinal scale) is due to an effective 'moist static stability' since heating in this model is proportional to vertical velocity. At the CISK-neutral point, latent heat release is exactly balanced by adiabatic cooling. This state of affairs is sometimes referred to as 'free-ride' balance (Fraedrich and McBride, 1989) and represents approximate convective neutrality of the tropical atmosphere. It is not limited to a two-level model, as latent heat release can be equal to, or some fixed fraction of, adiabatic cooling at several vertical levels.

We now inquire whether a similar reduction of phase speed occurs in a beta-plane model with conditional heating. The question is important because, in the real atmosphere, effective 'moist stability' is a function of the prevailing vertical motion and latent heat release. We could either assume some fixed zonal variation of stability representing the Walker circulation and underlying SST gradient, or make the stability solution-dependent, *i.e.*, with conditional heating in a uniform basic state. The second approach was adopted here although the first is equally relevant to the atmosphere.

For simplicity, solutions were sought with linearized evaporation-wind feedback and easterly mean flow in a two-layer model. Table 1 displays the growth rate and intrinsic phase speed of unstable Kelvin waves near the CISK-neutral point obtained from numerical integrations with $(M, N) = (128, 44)$ and $\nu = 2.5 \cdot 10^5\ m^2\,s^{-1}$. It is apparent that phase speeds decrease quite dramatically approaching the neutral point, and continue to decrease entering the CISK-unstable region. (The advective influence of low-level $\bar{u}$ was included in these simulations, but not in Table 1. For an easterly mean flow, ground-based phase speeds $c = \hat{c} + \bar{u}$ are less than the tabulated values of $\hat{c}$.) Similar results were recently obtained by Kirtman and Vernekar (1993) using a Kelvin wave model with unconditional heating, and by Xie *et al.* (1993). Horizontal structure in most

cases was that of a Kelvin wave, but with nonzero meridional velocity as $\hat{c} \to 0$. Table 1 suggests that the phase speed of the whole disturbance, including dry region, is controlled by an effective 'moist stability' of the wet region. Phase speeds were also observed to decrease slowly with decreasing A as expected from linear theory. The CISK-stable region near $A = 0$ was problematic since nonzero diffusion was always required for computational purposes and growth rates are expected to be very weak. We were therefore unable to verify whether, as expected, the Kelvin modes described in Table 1 are contiguous to the westward-propagating hybrid structure obtained for $\bar{u} > 0$.

3. Conclusions

The effect of conditional heating on unstable Kelvin waves generated by CISK and evaporation-wind feedback was investigated with multi-level models on the equatorial beta-plane. In agreement with previous analytic and numerical results (Crum and Dunkerton, 1992), the scale-selection catastrophe of linear heating is modified but not eliminated by conditional heating. There is only one region of upward motion, but it contracts to the smallest zonal scale allowed by diffusion or model resolution. The pathology is most severe in the CISK-unstable case. With evaporation-wind feedback, Kelvin-wave instability is possible apart from CISK; growth rate in this case increases slowly with decreasing diffusion, but contraction is still observed. We note incidentally that eastward moving evaporation-wind modes can be generated on the beta-plane with no mean wind, contrary to previous assumptions requiring an easterly background flow. In a westerly mean flow, a westward-propagating hybrid structure is obtained.

While evaporation-wind feedback is an improvement over CISK, as far as scale selection is concerned, the suggestion that unstable Kelvin waves explain the tropical intraseasonal oscillation is inadequate. The first and most obvious problem is that zonal phase speeds are too fast compared to the observed propagation of intraseasonal oscillations and superclusters. In multi-level models with specified vertical heating profile this is due to the coalescence or destabilization of moist stable modes having phase speeds similar to those of dry Kelvin modes. In a two-level model, on the other hand, destabilization appears possible at an arbitrarily small phase speed when the effective moist stability is reduced to zero and the parameter A controlling the strength of evaporation-wind feedback is small. This state of affairs represents approximate convective neutrality of the tropical atmosphere. In the limit of neutral stability, unstable waves on the beta-plane depart from a pure Kelvin structure as required by linear theory. While slow eastward propagation appears possible in this case, we are bothered by the sensitivity to parameters (q, A) and wonder how atmospheric conditions could ever conspire to produce a robust TIO with well-defined spectral peak.

Two additional problems are the failure to generate any significant subtropical structure or hierarchy of convection within active phases of the TIO. These observations suggest, respectively, the importance of latitudinal shear (Dunkerton, 1990) and group velocity of synoptic-scale waves (Liebmann and Hendon, 1990). The fact that moisture convergence and evaporation-wind feedback actually take place in a hierarchy of squall lines, mesoscale convective complexes, and synoptic waves also reveals the inadequacy of a cumulus parameterization that depends only on large-scale variables.

Acknowledgements

This research was supported by the National Science Foundation, Grants ATM-8819582, ATM-9123797, ATM-9213437, and by the Air Force Office of Scientific Research, Contracts F49620-89-C-0051 and F49620-92-C-0033. Numerical computations were performed at the National Center for Atmospheric Research, which is sponsored by the National Science Foundation.

References

Chang, C.-P. and H. Lim, 1988: Kelvin wave-CISK: a possible mechanism for the 30–60 day oscillations. *J. Atmos. Sci.*, **45**, 1709–1720.

Charney, J.G. and A. Eliassen, 1964: On the growth of the hurricane depression. *J. Atmos. Sci.*, **21**, 68–75.

Crum, F.X. and D.E. Stevens, 1983: A comparison of two cumulus parameterization schemes in a linear model of wave-CISK. *J. Atmos. Sci.*, **40**, 2671–2688.

Crum, F.X. and T.J. Dunkerton, 1992: Analytic and numerical models of wave-CISK with conditional heating. *J. Atmos. Sci.*, **49**, 1693–1708.

Dunkerton, T.J., 1990: Eigenvalues and horizontal structure of divergent barotropic instability originating in tropical latitudes. *J. Atmos. Sci.*, **47**, 1288–1301.

Emanuel, K.A., 1987: An air-sea interaction model of intraseasonal oscillations in the tropics. *J. Atmos. Sci.*, **44**, 2324–2340.

Fraedrich, K. and J.L. McBride, 1989: The physical mechanism of CISK and the free-ride balance. *J. Atmos. Sci.*, **46**, 2642–2648.

Gill, A.E., 1980: Some simple solutions for heat-induced tropical circulations. *Q. J. Roy. Meteor. Soc.*, **106**, 447–462.

Hayashi, Y.-Y. and A. Sumi, 1986: The 30–40 day oscillations simulated in an 'aquaplanet' model. *J. Meteor. Soc. Japan*, **64**, 451–467.

Houghton, J.T., 1986: *The Physics of Atmospheres*. Second edition. Cambridge University Press. 271 pp.

Kirtman, B. and A. Vernekar, 1993: On wave-CISK and the evaporation-wind feedback for the Madden-Julian oscillation. *J. Atmos. Sci.*, **50**, 2811–2814.

Knutson, T.R. and K.M. Weickmann, 1987: 30–60 day atmospheric oscillation: composite life cycles of convection and circulation anomalies. *Mon. Wea. Rev.*, **115**, 1407–1436.

Lau, K.-M. and L. Peng, 1987: Origin of low frequency (intraseasonal) oscillations in the tropical atmosphere. Part I: The basic theory. *J. Atmos. Sci.*, **44**, 950–972.

Lau, N.-C., I.M. Held and J.D. Neelin, 1988: The Madden-Julian oscillation in an idealized general circulation model. *J. Atmos. Sci.*, **45**, 3810–3832.

Liebmann, B. and H.H. Hendon, 1990: Synoptic-scale disturbances near the equator. *J. Atmos. Sci.*, **47**, 1463–1479.

Madden, R.A. and P.R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708.

Madden, R.A. and P.R. Julian, 1972: Description of global scale circulation cells in the tropics with a 40–50 day period. *J. Atmos. Sci.*, **29**, 1109–1123.

Neelin, J.D., I.M. Held and K.H. Cook, 1987: Evaporation-wind feedback and low frequency variabilty in the tropical atmosphere. *J. Atmos. Sci.*, **44**, 2341–2348.

Rui, H. and B. Wang, 1990: Development characteristics and dynamic structure of tropical intraseasonal convection anomalies. *J. Atmos. Sci.*, **47**, 357–379.

Swinbank, R., T.N. Palmer and M.K. Davey, 1987: Numerical simulations of the Madden and Julian oscillation. *J. Atmos. Sci.*, **45**, 774–788.

Xie, S.-P., A. Kubokawa and K. Hanawa, 1993: Nonlinear evaporation-wind feedback mode. Preprint Volume, Ninth Conference on Atmospheric and Oceanic Waves and Stability, San Antonio, Texas, 385–386.

Yano, J.-I. and K. Emanuel, 1991: An improved model of the equatorial troposphere and its coupling with the stratosphere. *J. Atmos. Sci.*, **48**, 377–389.

Yoshizaki, M., 1991: Selective amplification of the eastward-propagating mode in a positive-only wave-CISK model on an equatorial beta plane. *J. Meteor. Soc. Japan*, **69**, 353–373.

赤道ベータ面における条件付き加熱を持つ CISK と蒸発–風フィードバック

Francis X. Crum and Timothy J. Dunkerton

(Northwest Research Associates)

赤道ベータ面の多層モデルを用いて、CISK と蒸発–風フィードバックによって発現する不安定なケルビン波への条件付き加熱の影響を調べた。今までの解析的な結果と同様に、(条件付きでない) 線型的な加熱の場合に (無限小のスケールが卓越する) スケール選択が破綻する問題は、条件付き加熱によって緩和されるものの解消はされない。条件付き加熱の上昇流域は一つであり、その水平スケールは粘性やモデルの解像度によって決まる最小のスケールになる。この傾向は CISK による不安定な場合にもっとも顕著である。蒸発–風フィードバックによってケルビン波の不安定は CISK と無関係に起こる。この場合、増幅率は粘性が小さくなるにつれゆるやかに大きくなるが、水平スケールはやはり小さくなる。今までは東風が必要と考えられていた東進する蒸発–風フィードバック不安定モードが、一般風のない場合にもあらわれるのが示された。

スケール選択に関しては蒸発–風フィードバックの方が CISK よりも良い。しかし、(1) 水蒸気収束フィードバックの強さを決めるパラメータ q が CISK 中立点に近い値を持ち、(2) 蒸発–風フィードバックの強さを決めるパラメータ A が小さく、(3) 加熱分布は鉛直流に比例する場合以外には、東西方向の位相速度は熱帯季節内振動やスーパークラスターのものに比べて非常に大きい。こうした三つの条件を満たす場合には、2 層モデルで小さな正の位相速度が得られる。

Horizontal Buoyancy Flux of Internal Gravity Waves in Vertical Shear

Timothy J. Dunkerton

Horizontal Buoyancy Flux of Internal Gravity Waves in Vertical Shear

By Timothy J. Dunkerton

Northwest Research Associates, P.O. Box 3027, Bellevue, WA 98009, U.S.A.

(Manuscript received 31 October 1994, in revised form 18 April 1995)

Abstract

The equations of motion describing two-dimensional internal gravity waves were analyzed to derive an expression for the horizontal buoyancy flux—*i.e.*, the correlation between buoyancy (or temperature) and horizontal velocity—in various cases involving vertical shear. It is shown that, although this correlation vanishes at zeroth order for a single wave, its first order contribution is nonzero due to shear, even for steady, conservative, incompressible waves. Departures from steady, conservative motion or quasi-compressibility also cause a nonzero correlation. Several cases were analyzed and some numerical results obtained for waves approaching a critical layer of reduced intrinsic phase speed. With weak shear, the buoyancy flux is small relative to vertical momentum flux, as expected from the perturbation theory. Strong vertical shear enhances the buoyancy flux within the shear zone and causes partial reflection beneath, producing a nonzero correlation (at zeroth order) in this region. These effects may explain recent observations of zonal wind and temperature cospectra in the equatorial lower stratosphere.

1. Introduction

Gravity waves are important for the transport of momentum, heat and constituents in the middle atmosphere. Their vertical, rather than horizontal, transport is emphasized in the literature since the waves grow with height and are able, for example, to transport a relatively small amount of momentum from the lower to upper atmosphere, causing significant mean flow acceleration in the mesosphere (Lindzen, 1981; Fritts, 1984). Little attention has been paid to the horizontal transport of momentum by gravity waves, and even less to the horizontal transport of heat which is formally smaller (by an order of magnitude) in WKB theory.

Whether observations assign a significant role to the horizontal transport is uncertain, but the interpretation of observations may require an understanding of horizontal fluxes. Recently Sato *et al.* (1994) discovered that zonal wind (u') and temperature (T') fluctuations in the equatorial lower stratosphere, with period ~1–3 days, display a remarkable correlation that changes sign according to the phase of the quasi-biennial oscillation (QBO). Equatorial Kelvin waves were previously identified from the quadrature spectrum $Q(u', T'; \omega)$ which has a significant peak at ~5–20 days in the equatorial lower stratosphere (Wallace and Kousky, 1968; Hirota, 1978; Salby *et al.*, 1984; Maruyama, 1991). Recent observations suggest an important contribution to the momentum flux from disturbances with periods below 5 days (Pfister *et al.*, 1993; Maruyama, 1994; Bergman and Salby, 1994). Noting the peculiar behavior of the cospectrum in their data, Sato *et al.* (1994) question whether the observed phase relation between u' and T' is consistent with Kelvin or internal gravity waves.

Sato *et al.* were mainly concerned with the interpretation of short-period disturbances, but it is obvious from their analysis that a similar problem arises with respect to ~5–20 day disturbances traditionally identified as Kelvin waves. At Singapore, the integrated ~5–20 day quadrature $Q(u', T')$ and cospectrum $C(u', T')$ power are greatly enhanced during the onset of the QBO westerly phase at each level.[1] Their ratio is typically 1.5:1. The spikes of ~5–20 day quadrature power are consistent with those of eastward-propagating waves encountering a critical layer in descending QBO westerlies. Meridional velocity fluctuations (v') are uncorrelated (or negligible), supporting a Kelvin-wave interpretation. But the cospectrum of these waves, and that of short-period disturbances, requires explanation.

The purpose of this paper is to discuss the relation between horizontal wind and temperature (or, equivalently, buoyancy) for two-dimensional internal

©1995, Meteorological Society of Japan

1 As shown in Fig. 10 of Sato *et al.* (1994). Spectral analysis of Singapore rawinsonde data by the author supports their conclusions.

gravity waves in vertical shear. It is shown theoretically and numerically that the correlation $\overline{u'T'}$ is nonzero and proportional to the shear—even for steady, conservative, incompressible waves. For vertical shears typical of the QBO, the effect of compressibility is small, other than causing the waves to grow with height. Realistic damping likewise does not significantly alter the effect of shear. Numerical solutions demonstrate that superposition of upward and downward propagating waves creates a significant horizontal buoyancy flux beneath the shear zone, consistent with zeroth order wave solutions. Finally, we note that although the horizontal flux is formally nondivergent in the situation examined here, real wavepackets are of finite extent such that the horizontal transport of buoyancy is nonzero.

2. Theoretical results

The equations of motion for two-dimensional, nonrotating, hydrostatic, quasi-compressible, linear perturbations on a zonal-mean flow are

$$D_t u' + w'\bar{u}_z + \phi'_x = X' \tag{2.1a}$$

$$D_t \phi'_z + w'N^2 = Q' \tag{2.1b}$$

$$u'_x + \frac{1}{\rho_0}(\rho_0 w')_z = 0 \tag{2.1c}$$

in the notation of Holton (1975), where u, w are zonal and vertical velocity, ϕ is geopotential, $\rho_0 = \rho_s \exp(-z/H)$ is basic state density, and X, Q represent sources or sinks of momentum and heat. Primed quantities are deviations from the zonal mean, denoted by an overbar; subscripts x, z indicate differentiation, and

$$D_t \equiv \frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}. \tag{2.2}$$

a. Wave action law

Following Andrews and McIntyre (1976), parcel displacements and auxiliary variables may be defined as

$$D_t \xi' = u^\ell \tag{2.3a}$$

$$D_t \zeta' = w' \tag{2.3b}$$

$$D_t q' = Q' \tag{2.3c}$$

where

$$u^\ell = u' + \zeta'\bar{u}_z. \tag{2.4}$$

Assuming that zonally averaged quantities are temporally constant to leading order in wave amplitude, (2.1a–c) may be written

$$D_t u^\ell + \phi'_x = X' \tag{2.5a}$$

$$\phi'_z + \zeta' N^2 = q' \tag{2.5b}$$

$$\xi'_x + \frac{1}{\rho_0}(\rho_0 \zeta')_z = 0. \tag{2.5c}$$

Multiplication of (2.5a,b) by ξ'_x, ζ'_x gives

$$(\overline{\xi'_x u^\ell})_t + (\overline{\xi'_x \phi'_x} + \overline{\zeta'_x \phi'_z}) = \overline{\xi'_x X'} + \overline{\zeta'_x q'} \tag{2.6}$$

noting that

$$\begin{aligned}\overline{\xi'_x D_t u^\ell} &= (\overline{\xi'_x u^\ell})_t - \overline{D_t \xi'_x u^\ell} \\ &= (\overline{\xi'_x u^\ell})_t - \overline{u^\ell_x u^\ell} = (\overline{\xi'_x u^\ell})_t\end{aligned}$$

and $\overline{\zeta'_x \zeta'} = 0$. Substitution of (2.5c) in (2.6) implies that

$$(\overline{\xi'_x u^\ell})_t + \frac{1}{\rho_0}(\overline{\rho_0 \zeta'_x \phi'})_z = \overline{\xi'_x X'} + \overline{\zeta'_x q'}. \tag{2.7}$$

This equation describes conservation of wave action, a quadratic (second-order) quantity in wave amplitude. A three-dimensional version of this equation (in y, z, t) was derived by Andrews and McIntyre(1976):

$$\frac{\partial A}{\partial t} + \frac{1}{\rho_0}\nabla \cdot \mathbf{B} = \mathcal{F} \tag{2.8}$$

where A is wave action density, $\mathbf{B}$ is wave action flux, and $\mathcal{F}$ represents the departure from conservative motion. For steady, conservative waves, $\mathbf{B}$ is nondivergent. The famous theorem of Eliassen and Palm (1961), extended by Andrews and McIntyre, relates the mean flow tendency $\bar{u}_t$ to the divergence of $\mathbf{B}$ and other departures from steady, conservative motion.

The vertical flux in (2.7) is akin to a 'radiation stress' describing the impulse to an overlying layer of fluid by pressure forces ϕ' averaged over a wave-distorted fluid interface $z + \zeta'(x, z, t)$. (Refer to Andrews *et al.*, 1987, p. 325 for a diagram.) This stress is nonzero, since the quantities ζ'_x, ϕ' are in phase for internal gravity waves, but it does not vary in height if the waves are steady and conservative.

The radiation stress may be related to the vertical flux of horizontal momentum as

$$\begin{aligned}\overline{\zeta'_x \phi'} &= -\overline{\zeta' \phi'_x} \\ &= \overline{\zeta'(D_t u^\ell - X')} \\ &= (\overline{\zeta' u'})_t - \overline{u'w'} + \frac{1}{2}(\overline{\zeta'^2})_t \bar{u}_z - \overline{\zeta' X'}\end{aligned} \tag{2.9}$$

from which it is apparent that the momentum flux differs from the radiation stress only on account of departures from steady, conservative motion. In general, the components of Eliassen-Palm flux satisfy this type of relationship which—together with the wave action law (2.7) and mean tendency equations—leads to a 'nonacceleration' theorem (Andrews and McIntyre, 1976).

b. Correlation of zonal wind and temperature

For hydrostatic waves the temperature perturbation is proportional to ϕ'_z (Holton, 1975). The correlation of zonal wind and temperature is therefore proportional to

$$\rho_0 \overline{u'\phi'_z} = \rho_0 \overline{u'(q' - \zeta' N^2)}$$

$$= \rho_0 \overline{u'q'} - \frac{\rho_0 N^2}{k^2} \overline{u'_x \zeta'_x}$$

$$= \rho_0 \overline{u'q'} - \frac{N^2}{\rho_0 k^2 \hat{c}} \overline{(\rho_0 w')_z \rho_0 w'} \qquad (2.10)$$

for each harmonic $\exp i(kx - \omega t)$, where k is zonal wavenumber, ω is frequency, $c = \omega/k$ is phase speed, and $\hat{c} = c - \bar{u}$ is intrinsic phase speed. (Use of the harmonic form is convenient, but unnecessary to the argument as shown at the end of this subsection.) The continuity equation (2.1c) was inserted on the last line of (2.10), together with the vertical displacement (2.3b), *i.e.*,

$$-\hat{c}\zeta'_x = w' \qquad (2.11)$$

for steady waves.

A similar expression for $\rho_0 \overline{u'\phi'_z}$ may be derived using the streamfunction

$$-\rho_0 u' = (\rho_0 \psi')_z \qquad (2.12a)$$

$$\rho_0 w' = (\rho_0 \psi')_x \qquad (2.12b)$$

From (2.11), $-\hat{c}\zeta' = \psi'$ so that

$$\rho_0 \overline{u'\phi'_z} = \rho_0 \overline{u'(q' - \zeta' N^2)}$$

$$= \rho_0 \overline{u'q'} - \frac{N^2}{\rho_0 \hat{c}} \overline{(\rho_0 \psi')_z \rho_0 \psi'} \qquad (2.13)$$

which agrees with (2.10) on account of the harmonic form assumed in the derivation.

Apart from a thermal damping term, the correlation of zonal wind and temperature evidently depends on the vertical variation of wave amplitude squared, as determined by $\rho_0 \psi'$ or $\rho_0 w'$:

$$\rho_0 \overline{u'\phi'_z} = -\frac{N^2}{2\rho_0 k^2 \hat{c}} (\rho_0^2 \overline{w'^2})_z + O(q') \qquad (2.14)$$

Consider now the special case of a steady, conservative upward-propagating wave in a basic state $(\rho_0, \bar{u}, N)$ independent of height, for which

$$w' = -\frac{ku'}{m} = \frac{k\hat{c}}{N} u' \qquad (2.15)$$

where m is vertical wavenumber. Substituting for a factor w' in (2.14),

$$\rho_0 \overline{u'\phi'_z} = -\frac{N^2}{2\rho_0 k^2 \hat{c}} (\rho_0^2 \overline{u'w'} k\hat{c}/N)_z \qquad (2.16)$$

The quantity $\rho_0 \overline{u'w'}$ is independent of height by the wave action law (2.7) and (2.9), as are $\rho_0, k, \hat{c}$, and N, so the correlation $\rho_0 \overline{u'\phi'_z}$ vanishes identically. This is obvious from the perfect correlation (or anticorrelation) of u' and ϕ' to begin with, and the fact that ϕ'_z is exactly in quadrature with ϕ'.

Another special case involves the superposition of identical upward and downward waves with unequal amplitude:

$$\psi' = \mathrm{Re}[\psi_\uparrow \exp i(kx + mz - \omega t)$$

$$+ \psi_\downarrow \exp i(kx - mz - \omega t)]$$

$$= \mathrm{Re}\psi_\uparrow [\exp i(kx + mz - \omega t)$$

$$+ R \exp i(kx - mz - \omega t + \delta)] \qquad (2.17a)$$

$$\overline{\psi'^2} = \frac{1}{2} |\psi_\uparrow|^2 \left[1 + R^2 + 2R\cos(2mz - \delta)\right] \qquad (2.17b)$$

The correlation $\rho_0 \overline{u'\phi'_z}$ oscillates in the vertical about zero, and if $R = 1$, the momentum flux $\overline{u'w'}$ vanishes. Perfect reflection is analogous to a zonally propagating circulation cell in the $x - z$ plane confined between two rigid plates, such that vertical velocity vanishes on the boundaries. Suppose the propagation is eastward, and the motion adiabatic. Behind the upwelling there is cold air, due to adiabatic cooling, which partially overlaps the eastward (westward) motion at lower (upper) levels required by continuity. Warm air is ahead of the upwelling, overlapping the westward (eastward) motion at lower (upper) levels. The correlation $\overline{u'T'}$ is therefore negative (positive) in the lower (upper) layer, and zero at the midpoint and boundaries. By contrast, equal amounts of positive and negative momentum flux cancel identically in each layer when the zonal average is taken.

The WKB approximation provides a third example:

$$\psi'(x, z, t) = \mathrm{Re} A(Z) \exp i\Phi(x, z, t) \qquad (2.18)$$

where $k = \partial\Phi/\partial x$, $m = \partial\Phi/\partial z$, $\omega = -\partial\Phi/\partial t$, and $Z = \mu z$ is a 'slow' vertical coordinate such that A, m vary slowly in height, *i.e.*, as functions of Z. Then

$$\overline{\psi'_x w'_x} = \frac{k}{m} \left[\overline{\psi'_x \psi'_z} - \frac{1}{2}\mu k \ \mathrm{Im}(A^* A')\right]. \qquad (2.19)$$

This is analogous to the first case, but using a local value of m, together with $O(\mu)$ correction that varies slowly in height. If departures from steady, conservative motion are $O(\mu^2)$ or less, the momentum flux is invariant to this order, and a useful result is

$$\rho_0 \overline{u'\phi'_z} = -\frac{N^2}{2\rho_0 k^2 \hat{c}} (\rho_0 \overline{u'w'}) \cdot (\rho_0 k\hat{c}/N)_z + O(\mu^2) \qquad (2.20)$$

In (2.18–20) we have also assumed that the basic state density variation is $O(\mu^2)$, *i.e.*, that the mean flow varies rapidly over a density scale height, as observed in the QBO. Supposing for consistency that variations of N are likewise small, (2.20) reduces to

$$\rho_0 \overline{u'\phi'_z} = \frac{N}{2k\hat{c}} \rho_0 \overline{u'w'} \bar{u}_z + O(\mu^2). \qquad (2.21)$$

The correlation between zonal wind and temperature is proportional to vertical shear, an order of magnitude smaller (in μ) than the momentum flux.

For upward waves, the ratio of momentum flux to intrinsic phase speed is positive, so the sign of $\rho_0\overline{u'\phi'_z}$ in this case is determined by the vertical shear. Upward propagating waves with phase speeds of opposite sign will contribute in the same way to the correlation of zonal wind and temperature, but their momentum fluxes will tend to cancel. For downward waves, the sign of the correlation is unchanged since the sign of m in (2.15), and of $\overline{u'w'}/\hat{c}$, are both reversed.

Under the assumptions leading to (2.21) the same result may be obtained from the WKB approximation noting that $|\psi'| \propto |m|^{-1/2}$, substituting the value of m from (2.15) together with this proportionality into (2.14). If the shear is positive, $\overline{\psi'^2}$ decreases (increases) with height for waves with positive (negative) intrinsic phase speed; the correlation $\rho_0\overline{u'T'}$ is therefore positive. If the shear is negative, $\overline{\psi'^2}$ increases (decreases) with height for waves with positive (negative) intrinsic phase speed; the correlation is then negative. It is assumed that the sign of $\hat{c}$ is invariant, *i.e.*, that waves cannot propagate through critical levels.

If the unsteadiness, damping and density variation are $O(\mu)$, a more complicated expression for the single upward wave may be obtained from (2.14) by writing

$$\overline{w'^2} = \overline{u'w'}k\hat{c}/N + O(\mu) \tag{2.22}$$

so that

$$\begin{aligned}\rho_0\overline{u'\phi'_z} &= -\frac{N^2}{2\rho_0 k^2\hat{c}}\left(\rho_0^2\overline{u'w'}k\hat{c}/N\right)_z \\ &\quad +\rho_0\overline{u'q'} + O(\mu^2) \\ &= -\frac{N^2}{2k\hat{c}}\overline{u'w'}(\rho_0\hat{c}/N)_z \\ &\quad -\frac{N}{2k}(\rho_0\overline{u'w'})_z \\ &\quad +\rho_0\overline{u'q'} + O(\mu^2)\end{aligned} \tag{2.23}$$

The first term represents variations of the basic state at $O(\mu)$; the second term contains all the departures from steady, conservative motion at $O(\mu)$ implied by (2.7) and (2.9).

Three simple examples demonstrate the effects of density variation, damping, and saturation at $O(\mu)$: 1) If the waves are steady and conservative, and $\bar{u}$, N are constant in height (but not ρ_0), then from (2.14,15)

$$\begin{aligned}\rho_0\overline{u'\phi'_z} &= -\frac{N^2}{2\rho_0 k^2\hat{c}}\rho_0\overline{u'w'}\frac{k\hat{c}}{N}\rho_{0z} + O(\mu^2) \\ &= \frac{N}{2kH}\rho_0\overline{u'w'} + O(\mu^2)\end{aligned} \tag{2.24}$$

The incompressible value of m was substituted in this equation. 2) If the waves are steady but damped by mechanical and thermal relaxation, with equal coefficients $\alpha_M = \alpha_T = \frac{1}{2}\alpha$, and the basic state is again constant in height (including ρ_0), then it may shown from the wave action law (2.7) and zeroth order wave solutions that

$$B(z) = B_0 \exp -P(z) \tag{2.25a}$$

$$P(z) = \int_0^z D(z')dz' \tag{2.25b}$$

$$D(z) = \alpha N/k\hat{c}^2 \tag{2.25c}$$

where $B = \rho_0\overline{u'w'}$. In this case B decays with height, such that $B_z = -BP_z = -BD$, therefore

$$\begin{aligned}\rho_0\overline{u'\phi'_z} &= -\frac{N^2}{2\rho_0 k^2\hat{c}}\rho_0 B_z\frac{k\hat{c}}{N} + \rho_0\overline{u'q'} + O(\mu^2) \\ &= \frac{N}{2kH_D}\rho_0\overline{u'w'} + \rho_0\overline{u'q'} + O(\mu^2)\end{aligned} \tag{2.26}$$

where $H_D = k\hat{c}^2/N\alpha$ (Dunkerton, 1981). Note the similarity to the previous case, with H replaced by H_D. This formula could easily be extended to include $O(\mu)$ basic state variations. 3) Following Lindzen (1981), suppose that the waves are steady and conservative, growing with height until 'saturating' at some level, the criterion for saturation being that the waves are marginally unstable to convection. Above this level,

$$B = B_s = \rho_0\frac{k\hat{c}^3}{N} \tag{2.27}$$

so that

$$\begin{aligned}\rho_0\overline{u'\phi'_z} &= -\frac{N^2}{2\rho_0 k^2\hat{c}}(\rho_0^2 k^2\hat{c}^4/N^2)_z + \rho_0\overline{u'q'} \\ &\approx 2\rho_0\bar{u}_z\hat{c}^2 + \rho_0\overline{u'q'}.\end{aligned} \tag{2.28}$$

On the last line it was assumed for illustration that the mean flow varies rapidly over a density scale height. Note that the buoyancy flux at the breaking level, apart from the thermal damping term, is four times that implied by (2.21), but decreases rapidly above this level. More complete analysis of saturation may require consideration of the explicit forms of X', Q' that cause the waves to saturate in the first place.

The contribution from $\rho_0\overline{u'q'}$ may be evaluated for steady waves noting that $-\hat{c}q'_x = Q'$, or

$$k^2\hat{c}q' = -\alpha_T\phi'_{zx} = \alpha_T N^2\zeta'_x = -\alpha_T N^2 w'/\hat{c} \tag{2.29}$$

so that

$$\rho_0\overline{u'q'} = -\frac{\alpha_T N^2}{k^2\hat{c}^2}\rho_0\overline{u'w'}. \tag{2.30}$$

The sign is opposite to the momentum flux. The effect of thermal damping is to shift temperature anomalies forward in the direction of intrinsic phase propagation, producing a negative correlation $\overline{u'T'} < 0$ if $\overline{u'w'} > 0$.

Finally, a nondimensional measure of the horizontal buoyancy flux may be obtained from the ratio

$$\frac{k\overline{u'\sigma'_z}}{\overline{u'\sigma'_{zx}}} = \frac{k\hat{c}}{N^2}\frac{\overline{u'\sigma'_z}}{\overline{u'w'}} + O(\mu^2)$$
$$= \frac{\mathrm{Re}(\hat{\psi}^*\hat{w}_z)}{\mathrm{Im}(\hat{\psi}^*\hat{w}_z)} - \frac{\alpha_T}{k\hat{c}} + O(\mu^2) \qquad (2.31)$$

For a wave of harmonic form, this is equivalent to the ratio of cospectrum and quadrature power. In the three simple examples above, the respective nondimensional contributions are $(2mH)^{-1}$, $(2mH_D)^{-1}$, and $2\bar{u}_z/N$. In (2.21) the nondimensional value is $\bar{u}_z/2N$.

From the variety of cases analyzed here, it is obvious that a horizontal buoyancy flux can arise at $O(1)$ from wave reflection, or at $O(\mu)$ within a region of basic state variation, wave transience, saturation, or damping. The $O(\mu)$ contribution should not be downplayed, particularly if it is due to critical-level interaction within a zone of strong vertical shear. In this case, zonal-wind perturbations grow as $m^{1/2}$ up to the breaking level. The observed horizontal buoyancy flux might actually be larger within, rather than below, the shear zone on account of the factor of $\hat{c}$ appearing in the denominator of (2.21), or because of wave saturation (2.28). Saturation, however, would rapidly diminish its magnitude above the breaking level.

The harmonic form has been used for convenience, but is unnecessary. Starting from (2.4) the second term on the rhs of (2.10) may be expanded as follows:

$$\begin{aligned}
-\rho_0\overline{\zeta' u'} &= \rho_0\overline{\zeta'^2}\bar{u}_z - \rho_0\overline{\zeta' u^\ell} \\
&= \rho_0\overline{\zeta'^2}\bar{u}_z - \rho_0\overline{\zeta' D_t\xi'} \\
&= \rho_0\overline{\zeta'^2}\bar{u}_z - \rho_0(\overline{\xi'\zeta'})_t + \rho_0\overline{w'\xi'} \\
&= \rho_0\overline{\zeta'^2}\bar{u}_z - \rho_0(\overline{\xi'\zeta'})_t - \rho_0\overline{\psi'\xi'_x} \\
&= \rho_0\overline{\zeta'^2}\bar{u}_z - \rho_0(\overline{\xi'\zeta'})_t \\
&\quad + \frac{1}{\rho_0}\overline{(\rho_0\psi')(\rho_0\zeta')_z} \qquad (2.32)
\end{aligned}$$

Using (2.12a) it can be shown that

$$-2\rho_0\overline{\zeta' u'} = \rho_0\overline{\zeta'^2}\bar{u}_z - \rho_0(\overline{\xi'\zeta'})_t + \frac{1}{\rho_0}\frac{\partial}{\partial z}\left(\rho_0^2\overline{\psi'\zeta'}\right) \qquad (2.33)$$

Quadratic terms on the rhs are evaluated from the lowest order wave solutions. The last term contains the vertical momentum flux (proportional to $\rho_0 N\overline{\psi'\zeta'}$ for steady, conservative, slowly varying waves) and a basic state variation (ρ_0/N) inside the vertical derivative. Substitution of the harmonic form in this equation recovers (2.13).

c. Numerical solutions

The linear equations (2.1a–c) may be combined to give

$$\hat{\psi}'_{zz} + m^2\hat{w}' = 0 \qquad (2.34)$$

for a single harmonic, where

$$m^2 = \frac{N^2}{(c-\bar{u})^2} + \frac{\bar{u}_{zz}}{c-\bar{u}} \qquad (2.35)$$

for hydrostatic, incompressible, steady, conservative waves. The nonhydrostatic and quasicompressible versions may be found in Booker and Bretherton (1967) and Holton (1975), respectively. If mechanical and thermal damping coefficients are equal and independent of height, then $c \to c + ic_0$ where $c_0 = \alpha/k$. To illustrate the effect of damping, a small imaginary part may be added to the phase speed. Eq. (2.32) was solved using the tridiagonal algorithm (with 801 equally spaced gridpoints) subject to boundary conditions

$$\hat{\psi}(z_B) = 1 \qquad (2.36a)$$
$$\hat{\psi}_z(z_T) = im(z_T)\hat{\psi}(z_T) \qquad (2.36b)$$

choosing the sign of m in (2.34b) to correspond to an upward propagating wave. Mean flow profiles were of the form

$$\bar{u} = \bar{u}_B + (\bar{u}_T - \bar{u}_B)\cdot\left[1 + \tanh\left(\frac{z-z_M}{D}\right)\right]/2 \qquad (2.37)$$

and $N = .02\ \mathrm{s}^{-1}$. In the following, $z_B = 0$, $z_M = 7.5$ km, and $z_T = 15$ km. For case 1 (weak shear):

$$\begin{aligned}
u_B &= 0 \\
u_T &= 10\ \mathrm{ms}^{-1} \\
D &= 2\ \mathrm{ms}^{-1}\ \mathrm{km}^{-1} \\
c &= 15\ \mathrm{ms}^{-1}. \qquad (2.38)
\end{aligned}$$

For case 2 (strong shear):

$$\begin{aligned}
u_B &= -20\ \mathrm{ms}^{-1} \\
u_T &= 20\ \mathrm{ms}^{-1} \\
D &= 2\ \mathrm{ms}^{-1}\mathrm{km}^{-1} \\
c &= 25\ \mathrm{ms}^{-1}. \qquad (2.39)
\end{aligned}$$

These profiles are shown in Figs. 1a, 1c. Case 2 is representative of the QBO near 25 km, with maximum shear 10 $\mathrm{ms}^{-1}\mathrm{km}^{-1}$, corresponding to a Richardson number $Ri = 4$.

The horizontal buoyancy flux is shown in Figs. 1b, 1d for cases 1, 2 respectively, along with (vertically constant) momentum flux. The WKB approximation (2.21) is also shown (dashed line). In case 1, there is a small buoyancy flux in the shear zone, and none outside. Nor is there any evidence of reflection. The WKB solution is, for practical purposes, exact. The momentum flux is .00133366, almost identical to the no-shear value ($m = .001\bar{3}$). In case 2, there is a larger buoyancy flux in and beneath the shear zone, with wavy structure below due to interference

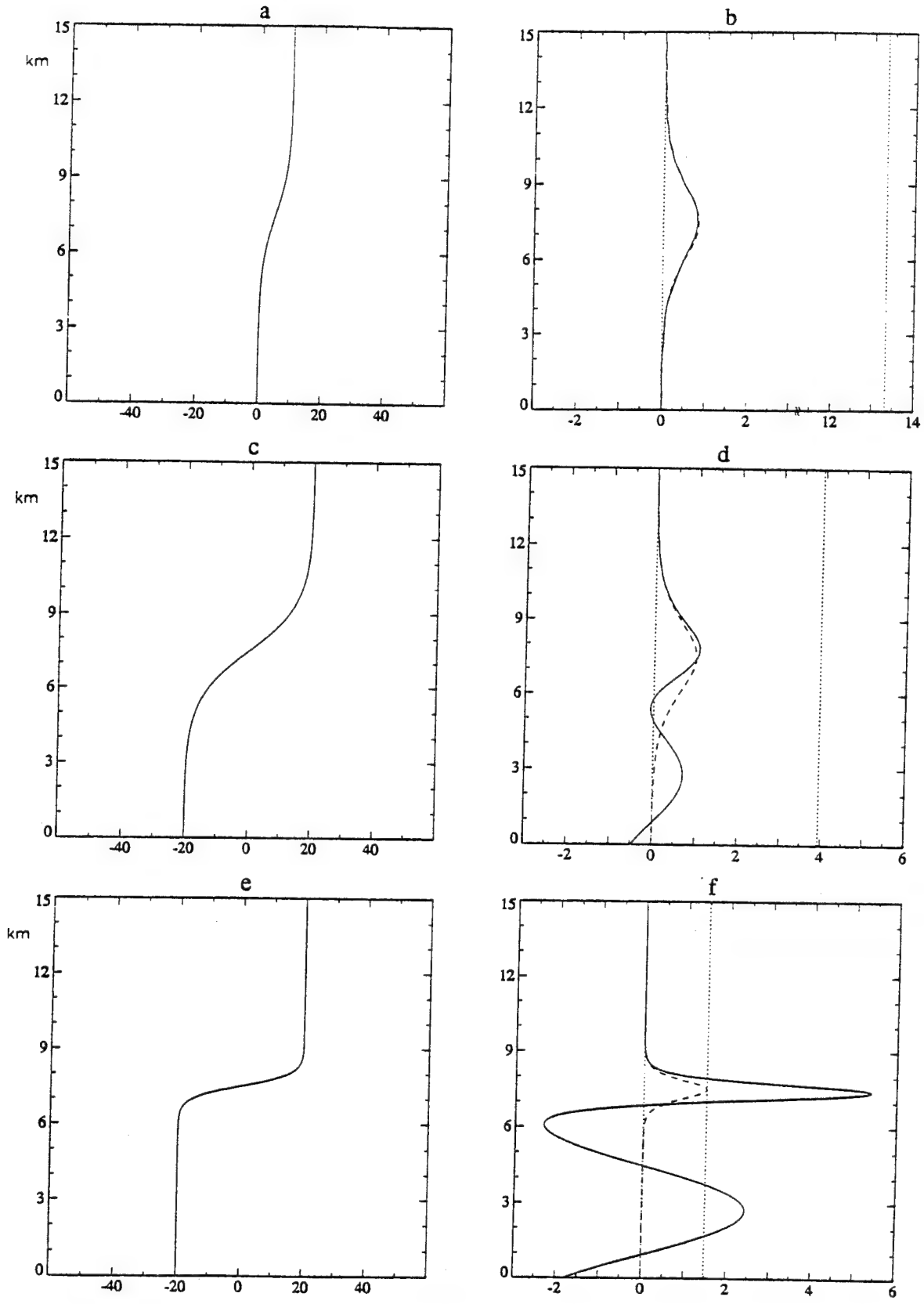


Fig. 1. Vertical profiles of mean zonal wind (units, ms^{-1}) for (a) case 1; (c) case 2; (e) variation of case 2. Horizontal buoyancy flux (units, m^{-1}) for (b) case 1; (d) case 2; (f) variation of case 2. Vertical flux of horizontal momentum (units, m^{-1}) is indicated by dotted line at right (constant in height). WKB solution for horizontal buoyancy flux is also shown (dashed line). Buoyancy (without $\hat{c}^{-1}$ factor) and momentum fluxes displayed in the figure correspond, respectively, to the real and imaginary parts of $-\psi^*\psi_z$.

of upward and downward waves in this region. The wavelength of the pattern is approximately half that of the wave itself (*cf.* 2.17). The WKB solution omits the small reflected component but is otherwise accurate. The momentum flux is .00039424, slightly less than the no-shear value ($m = .0004\bar{4}$).

Two variations of case 2 were examined. In the first, a small damping was added, $c_0 = 1$ ms^{-1}. The momentum flux now decays with height, markedly so as $\hat{c}$ diminishes entering the shear zone (not shown). The buoyancy flux is slightly more positive than before; the only major difference is that it is relatively larger when measured against the momentum flux in the shear zone. In the second variation (shown in Figs. 1e, 1f), $D = 0.5$ ms^{-1}km^{-1}; the shear is four times that of case 2 ($Ri \rightarrow 1/4$), with other parameters the same. The reflected wave is larger, and the momentum flux smaller (.000148514), so that the buoyancy flux is relatively much larger than in Fig. 1d. The ratio of horizontal buoyancy flux to vertical flux of horizontal momentum has no limit, since the momentum flux vanishes with perfect reflection. The amount of reflection observed in Fig. 1f is close to that predicted by Acheson's (1976) piecewise-continuous model ($R = 0.8$).

Other cases using the shear of Fig. 1e were analyzed for larger phase speeds ($|c| \gtrsim 50$ ms^{-1}) that might be associated with fast Kelvin or gravity waves and tides. Buoyancy fluxes are sharply peaked in the shear zone, with less evidence of reflection. When mean flows are reversed, the buoyancy fluxes reverse sign, but are similar in magnitude. When phase speeds are reversed, the buoyancy fluxes do not change sign and are again similar in magnitude.

d. Comments

In order to apply these examples to the atmosphere, wave amplitudes must be weighted by a factor $\exp(z/2H)$ and $O(\mu)$ corrections due to compressibility should be retained. If the major effect is wave growth, the cospectrum and quadrature power corresponding to that shown in the figures will have an exponential envelope, growing with height. Moreover, observable fields like u', T' will grow relative to ψ', as $\hat{c} \rightarrow 0$, by a factor $\hat{c}^{-1}$. These two effects will tend to emphasize features in the shear zone over those below. Apart from reflection, $\overline{u'T'}$ will maximize between the base of the shear zone and some point just above the breaking level (if one exists). A more detailed analysis of observations of Kelvin and gravity waves in relation to the descending QBO shear zones is warranted before drawing specific conclusions. Concerning Kelvin waves, it should be noted that their latitudinal scale contracts approaching the critical layer, causing wave amplitude to pile up near the equator more dramatically than in the two-dimensional problem. This is a WKB effect (with slowly varying approximation in time and height, but not latitude, as in Boyd, 1987) and does not require *vertical* reflection from sloping surfaces $\hat{\omega} = f$, although such reflection may be important if our two-dimensional examples are representative. Mean winds are sometimes asymmetric about the equator, so maximum wave amplitude may be shifted slightly off the equator as a result of reflection.

Time series of cospectrum power observed by Sato *et al.* (1994) reveal a qualitative difference between long and short-period disturbances at the equator. The cospectrum and quadrature power of ~5–20 day disturbances are positively enhanced at the onset of QBO westerlies, but are much smaller at other times. By contrast, the cospectrum power of short-period disturbances oscillates with the QBO between small positive and negative values of nearly equal magnitude. Although the quadrature power at short periods is markedly less than that of long-period waves, the relatively large magnitude and oscillatory behavior of the cospectrum $C(u', T')$ suggests that it may represent a real phenomenon. If so, two possible interpretations are:

1) Fast Kelvin waves, although transparent to QBO westerlies (Hirota, 1978; Salby *et al.*, 1984; Maruyama, 1991), might contribute to a small horizontal buoyancy flux as they encounter QBO shear zones. In order for $C(u', T')$ to be enhanced apart from $Q(u', T')$, the vertical shear must be extremely strong (*e.g.*, Fig. 1f).

2) The gravity wave spectrum contains eastward and westward propagating waves (Pfister *et al.*, 1993).[2] If waves of equal but opposite phase speed are incoherent, they will contribute alike to the horizontal buoyancy flux (2.21), but momentum fluxes will tend to cancel. The vertical shear need not be strong.

A possible problem with either of these explanations is that, unless phase speeds are large compared to the QBO, waves will be selectively amplified (as measured by u') in the QBO shear zones as their intrinsic phase speed is reduced. For these waves, it is difficult to see how $C(u', T')$ could be amplified without a similar (though possibly smaller) amplification of $Q(u', T')$. (Downward propagating waves would be attenuated by damping or saturation at higher levels, unless phase speeds are large.) Consistent with the role of vertical shear is the observation in Fig. 11 of Sato *et al.* (1994) that short-period $C(u', T')$ is amplified at slightly different altitudes depending on the phase of the QBO.

Given the likelihood that many gravity waves are present at the same time, it may be worthwhile to construct spectral analogues of the monochro-

2 The role of gravity waves in the QBO was emphasized by Dunkerton (1991) and Takahashi and Boville (1992) and is part of the original Lindzen and Holton (1968) theory.

matic formulae for $\overline{u'T'}$. One might assume, for example, that individual waves saturate just below critical levels, but that the resulting turbulence causes mechanical and thermal damping of other (non-breaking) waves. Presumably the waves most susceptible to damping in the westerly acceleration phase of the QBO would be those with phase speeds of positive sign and vice versa. Simultaneous observations of wind and temperature in the equatorial lower stratosphere, with better temporal resolution (*e.g.*, radar, lidar), would provide valuable information with which to evaluate these effects in the context of a gravity wave spectrum.

3. Conclusion

The equations of motion describing two-dimensional internal gravity waves were analyzed to derive an expression for the horizontal buoyancy flux—that is, the correlation between buoyancy (or temperature) and horizontal velocity—in various cases involving vertical shear. It was shown that, although this correlation vanishes at zeroth order (in μ) for a single wave, its first order contribution is nonzero due to shear, even for steady, conservative, incompressible waves. Departures from steady, conservative motion or quasi-compressibility also cause a nonzero correlation. Several cases were analyzed and some numerical results obtained for waves approaching a critical layer of reduced intrinsic phase speed. With weak shear, the buoyancy flux is small relative to momentum flux, as expected from the perturbation theory. Strong vertical shear enhances the buoyancy flux within the shear zone and causes partial reflection beneath, producing a nonzero $O(1)$ correlation in this region.

These effects may explain recent observations of zonal wind and temperature cospectra in the equatorial lower stratosphere (Sato *et al.*, 1994). Such observations, at least, do not immediately preclude their interpretation as Kelvin or gravity waves. Taking the analysis of this paper into account, the behavior of cospectrum and quadrature power relative to the QBO seems qualitatively consistent with such an interpretation for long and short-period disturbances, respectively. For long-period disturbances, equatorial $C(u', T')$ and $Q(u, t')$ are selectively enhanced during the onset of QBO westerly phases, but otherwise small, as expected for Kelvin waves. For short-period disturbances, $C(u', T')$ alternates sign with the QBO, as expected of a gravity wave spectrum containing both eastward and westward propagating components (Pfister *et al.*, 1993).

Although the theoretical results suggest that the cospectrum will be measurable, particularly in a critical layer, further quantitative analysis of the ratio C:Q is desirable as it may reveal useful information about the anisotropy of the spectrum or the nature of wave saturation and absorption. Simultaneous radar and lidar measurements in the equatorial lower stratosphere would be particularly valuable. A three-dimensional analysis of Kelvin waves in vertical shear and their associated horizontal buoyancy flux would significantly extend these results. Such topics are beyond the scope of this paper but fruitful avenues for exploration. A final point is that, although the horizontal buoyancy (or temperature) flux has been emphasized for its interpretive value—it is, after all, exactly nondivergent since independent of x—real wavepackets are of finite horizontal extent and may therefore contribute to a net horizontal transport under conditions more general than those considered here.

Acknowledgements

This research was supported by the Air Force Office of Scientific Research, Contract F49620-92-C-0033, and by the National Science Foundation, Grant ATM-9123797.

References

Acheson, D.J., 1976: On over-reflection. *J. Fluid Mech.*, **77**, 433–472.

Andrews, D.G. and M.E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–2048.

Andrews, D.G., J.R. Holton and C.B. Leovy, 1987: *Middle Atmosphere Dynamics.*, Academic Press, 489 pp.

Bergman, J.W. and M.L. Salby, 1994: Equatorial wave activity derived from fluctuations in observed convection. *J. Atmos. Sci.*, **51**, 3791–3806.

Booker, J.R. and F.P. Bretherton, 1967: The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.*, **27**, 513–539.

Boyd, J.P., 1987: The effects of latitudinal shear on equatorial waves, Part 1: theory and methods. *J. Atmos. Sci.*, **35**, 2236–2258.

Dunkerton, T.J., 1981: Wave transience in a compressible atmosphere, part 1: transient internal wave, mean-flow interaction. *J. Atmos. Sci.*, **38**, 281–297.

Dunkerton, T.J., 1991: Nonlinear propagation of zonal winds in an atmosphere with Newtonian cooling and equatorial wavedriving. *J. Atmos. Sci.*, **48**, 236–263.

Eliassen, A., and E. Palm, 1960: On the transfer of energy in stationary mountain waves. *Geophys. Publ.*, **22**, no. 3.

Fritts, D.C., 1984: Gravity wave saturation in the middle atmosphere: a review of theory and observations. *Rev. Geophys. and Space Phys.*, **22**, 275–308.

Hirota, I., 1978: Equatorial waves in the upper stratosphere and mesosphere in relation to the semi-annual oscillation of the zonal wind. *J. Atmos. Sci.*, **35**, 714–722.

Holton, J.R., 1975: *The Dynamic Meteorology of the Stratosphere and Mesosphere.* Amer. Meteor. Soc., 319 pp.

Lindzen, R.S., 1981: Turbulence and stress due to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86C**, 9707–9714.

Lindzen, R.S. and J.R. Holton, 1968: A theory of the quasi-biennial oscillation. *J. Atmos. Sci.*, **25**, 1095–1107.

Maruyama, T., 1991: Annual variations and QBO-synchronized variations of the equatorial wave intensity in the lower stratosphere at Singapore during 1961–1989. *J. Meteor. Soc. Japan*, **69**, 219–232.

Maruyama, T., 1994: Upward transport of westerly momentum due to disturbances of the equatorial lower stratosphere in the period range of about 2 days - a Singapore data analysis for 1983–1993. *J. Meteor. Soc. Japan*, **72**, 423–432.

Pfister, L., K.R. Chan, T.P. Bui, S. Bowen, M. Legg, B. Gary, K. Kelly, M. Proffitt and W. Starr, 1993: Gravity waves generated by a tropical cyclone during the STEP tropical field program: a case study. *J. Geophys. Res.*, **98**, 8611–8638.

Salby, M.L., D.L. Hartmann, P.L. Bailey and J.C. Gille, 1984: Evidence for equatorial Kelvin modes in Nimbus 7 LIMS. *J. Atmos. Sci.*, **41**, 220–235.

Sato, K., F. Hasegawa and I. Hirota, 1994: Short-period disturbances in the equatorial lower stratosphere. *J. Meteor. Soc. Japan*, **72**, 895–872.

Takahashi, M. and B.A. Boville, 1992: A three-dimensional simulation of the equatorial quasi-biennial oscillation. *J. Atmos. Sci.*, **49**, 1020–1035.

Wallace, J.M. and V.E. Kousky, 1968: Observational evidence of Kelvin waves in the tropical stratosphere. *J. Atmos. Sci.*, **25**, 900–907.

鉛直シアー流中での内部重力波の水平温度フラックス

Timothy J. Dunkerton

(Northwest Research Associates, U. S. A.)

2次元内部重力波の式を用いて、鉛直シアーも含めたいろいろな場合の水平温度フラックスの式を導いた。水平温度フラックスはゼロ次のオーダーではゼロとなり、定常で保存する非圧縮的な波であっても、シアーの効果により1次のオーダーでゼロではなくなる。定常でなく保存的でない波のとき、また準圧縮性の波の場合もゼロでなくなる。いくつかの場合を解析し、また波が臨界層に近付いた場合の数値的な結果も示す。弱いシアーのとき、温度フラックスは鉛直運動量フラックスに比べて小さい。強い鉛直シアーのとき温度フラックスはシアーとともに大きくなり、部分反射を生み出しゼロでない値をもつ。これらの効果は熱帯下部成層圏での東西風と温度のコスペクトルに関する最近の観測結果を説明するであろう。

Generation of Inertia-Gravity Waves in a Simulated Lifecycle of Baroclinic Instability

Donal O'Sullivan and Timothy J. Dunkerton

Generation of Inertia–Gravity Waves in a Simulated Lifecycle of Baroclinic Instability

DONAL O'SULLIVAN AND TIMOTHY J. DUNKERTON

Northwest Research Associates, Bellevue, Washington

(Manuscript received 3 June 1994, in final form 5 April 1995)

ABSTRACT

The excitation and propagation of inertia–gravity waves (IGWs) generated by an unstable baroclinic wave was examined with a high-resolution 3D nonlinear numerical model. IGWs arose spontaneously as the tropospheric jetstream was distorted by baroclinic instability and strong parcel accelerations took place, primarily in the jetstream exit region of the upper troposphere. Subsequent propagation of IGWs occurred in regions of strong windspeed—in the tropospheric and stratospheric jets, and in a cutoff low formed during the baroclinic lifecycle. IGWs on the flanks of these jets were rotated inward by differential advection and subsequently absorbed by the model's hyperdiffusion. Although absorption of IGWs at the sidewalls of the jet is an artifact of the model, IGW propagation was for the most part confined to regions with an intrinsic period shorter than the local inertial period. Only a few IGWs were able to penetrate the middle stratosphere, due to weak winds or an unfavorable alignment of wavevector with respect to the mean flow.

IGWs are important both as a synoptic signal in the jetstream, which may influence subsequent tropospheric developments, and as a source of isentropic or cross-isentropic mixing in the lower stratosphere. The authors' results demonstrated for the first time numerically a significant isentropic displacement of potential vorticity isopleths due to IGWs above the tropopause. Since conditions for IGW propagation are favorable within a jet, a region of strong isentropic potential vorticity gradient, it is likely that inertia–gravity waves affect the permeability of the lower stratospheric vortex and may in some instances lead to stratosphere–troposphere exchange.

1. Introduction

Inertia–gravity waves (IGWs) are ubiquitous within the atmosphere, especially the upper troposphere and lower stratosphere. These motions are characterized by short vertical wavelengths (~1–5 km) and small intrinsic frequencies $\hat{\omega}$ near the Coriolis frequency f (Barat 1983; Yamanaka and Tanaka 1984; Thomas et al. 1992; Sato 1994). Their horizontal scale, which is usually estimated rather than measured, is thought to be large (~200–1000 km) compared to mesoscale gravity waves. While momentum transport is ascribed to mesoscale waves (Fritts 1984), it is the role of IGWs in constituent mixing, and not momentum transport, that is probably more relevant. IGW mixing may be isentropic (Pierce and Fairlie 1993; Ledwell et al. 1993), like that of synoptic-scale baroclinic waves, or cross isentropic due to instabilities arising within the wave field (Dunkerton 1984; Fritts and Rastogi 1985). Notwithstanding the importance of IGW propagation, breaking, and absorption, one would also like to know something about the excitation of these waves and whether their geographical and climatological variability can be understood in terms of the postulated forcing mechanisms.

The origin of inertia–gravity waves is ambiguous because regions of strong mean flow are a favored locus of IGW excitation *and* propagation (Dunkerton 1984). IGWs observed within or above the tropospheric jetstream were not necessarily excited there: due to their slow propagation, the source may have been as far away as the Tropics (Ushimaru and Tanaka 1990). Case studies often suggest a connection between IGWs and nearby synoptic weather systems (Sato 1989, 1993, 1994; Chan et al. 1991; Thomas et al. 1992; Eckermann and Vincent 1993). Therefore it is likely that extratropical sources dominate. That gravity wave variance increases above oceanic frontal systems and jetstreams (Fritts and Nastrom 1992) indicates that sources other than topography are important—an essential observation for understanding gravity wave transport in the Southern Hemisphere and Tropics. Part of the increase is undoubtedly due to high-frequency waves launched by moist convection or shear instabilities within frontal zones. (It was primarily mesoscale variance that was observed by Fritts and Nastrom.) However, one also expects a low-frequency IGW component due to the unstable baroclinic wave that would exist in the absence of convection, arising from the transiency of tropospheric flow as it relaxes to a balanced state. In a rotating fluid this process is referred to as geostrophic adjustment (Blumen 1972; Luo and Fritts 1993, and

Corresponding author address: Dr. Donal O'Sullivan, North West Research Associates, Inc., P.O. Box 3027, Bellevue, WA 98009-3027.

© 1995 American Meteorological Society

/ams v3628 1203 Mp 1 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203

references therein). During the adjustment, gravitational oscillations emanate and disperse away from the source. Geostrophic adjustment was interpreted by Vallis (1992) as a process that minimizes the energy for a fixed distribution of potential vorticity (PV). This interpretation assumes that the adjustment is rapid compared to the evolution of PV. General methods for finding energy extrema that allow rearrangements of PV as a materially conserved quantity were discussed by Vallis et al. (1989) and Shepherd (1990).

Many studies examined geostrophic adjustment in an idealized way, as linear relaxation from an initial imbalance of mass or momentum, with simplified (Cartesian or axisymmetric) geometry. In real life the adjustment involves complicated, three-dimensional motion with short vertical and (ultimately) horizontal scales. The need for consistent vertical resolution in primitive equation models was stressed by Lindzen and Fox-Rabinovitz (1989), and their comments pertain equally well to geostrophic adjustment. Only within the last few years have computational resources become available to make a realistic modeling study of cyclostrophic or geostrophic adjustment practical (e.g., Lelong and McWilliams 1993). One rather fundamental point that has not been extensively addressed is the mechanism by which an initially balanced flow evolves into an unbalanced state, followed by geostrophic adjustment (Van Tuyl and Young 1982; Duffy 1990). It is expected that inertia–gravity waves arise spontaneously in primitive equation models (Warn and Menard 1986). Since IGWs are problematic to weather prediction, the emphasis has been mainly on how to eliminate them, rather than explicit modeling of IGW propagation and absorption. This omission is surprising in view of the large body of literature on ageostrophic circulations within synoptic weather systems (e.g., Cammas and Ramond 1989; Keyser et al. 1989, and references therein) and their role in frontogenesis, tropopause folding, and stratosphere–troposphere exchange (Shapiro 1980; Keyser and Shapiro 1986; Shapiro and Keyser 1990). By analogy to tidal theory, the ageostrophic motion contains a "forced" evanescent component (the meridional circulation) and a "free" radiating component (the inertia–gravity waves).

Recently the authors performed numerical simulations to examine the evolution of IGWs in a lifecycle of midlatitude baroclinic instability. Patterned after the classic experiments of Simmons and Hoskins (1978, 1980), our simulations described the nonlinear lifecycle of an unstable baroclinic wave assuming wavenumber 6 symmetry on the sphere but incorporated much higher vertical resolution and a domain extending into the lower stratosphere in order to explicitly model the evolving inertia–gravity waves and their radiation into the middle atmosphere. This work was complementary to several other studies (beginning with Ley and Peltier 1978; see also Uccelini and Koch 1987; Gall et al. 1988; Garner 1989; Benard et al. 1992a,b; Blumen 1992; Snyder et al. 1993) concerned with gravity wave excitation in relation to frontogenesis. We attempted to simulate inertia–gravity waves excited in the upper troposphere by an evolving baroclinic instability, considering their vertical as well as horizontal propagation. Numerical results demonstrated the role of baroclinic instability in exciting inertia–gravity waves, due to rapid deformation of upper-tropospheric flow and strong parcel accelerations.

Although our integrations were designed to yield well-resolved inertia–gravity waves independent of model resolution, only the excitation and initial propagation of IGWs could be realistically simulated. Their asymptotic propagation and breakdown requires vertical and horizontal resolution well beyond that used here (e.g., Dunkerton and Robins 1992). Resolved patterns of horizontal velocity divergence were nonetheless quite complex, containing an assortment of gravity waves and IGWs due to surface frontogenesis as well as parcel accelerations in the upper-level jet. This paper emphasizes geostrophic adjustment as the numerically best resolved, and probably most fundamental, mechanism of IGW excitation. Section 2 describes the numerical model. Section 3 discusses the baroclinic lifecycle and the resulting IGWs, examining IGW wave parameters and Lagrangian parcel accelerations. Some IGW-related issues are discussed briefly in section 4.

2. Model description

Simulations were performed with a slightly modified version of the 3D, global, hydrostatic primitive equation model of Young and Villere (1985). This is a spectral transform model with finite differencing in the vertical. There is no topography or moisture. Subgrid-scale processes are represented by sixth-order horizontal hyperdiffusion $\nu_6 \nabla_H^6$ acting on the vorticity, divergence, and temperature fields. The spectral truncation in runs described here, unless stated otherwise, was triangular at total horizontal wavenumber 126 (denoted T126)—approximately equivalent to 1° of horizontal resolution. Hyperdiffusion is needed to prevent nonphysical behavior as sharp horizontal gradients are generated at the smallest resolvable scales. In practice, ν_6 corresponding to an e-folding time (τ) of 1 h or less for the smallest scale is sufficient for stable integrations. At T126, $\nu_6 = 1.36 \times 10^{25}$ m^6 s^{-1} was used, corresponding to a damping time of about 20 min for the smallest resolved scale.

The vertical coordinate was uniformly spaced in log pressure with $\Delta z = 700$ m, unless stated otherwise, and the domain extended vertically from $z = 0$–35 km (51 levels). Boundary conditions assumed zero geometric vertical velocity at the lowest level, while log pressure vertical velocity was zero at the uppermost level. There was no surface friction.

As discussed in section 3d, comparison simulations were performed at various combinations of horizontal

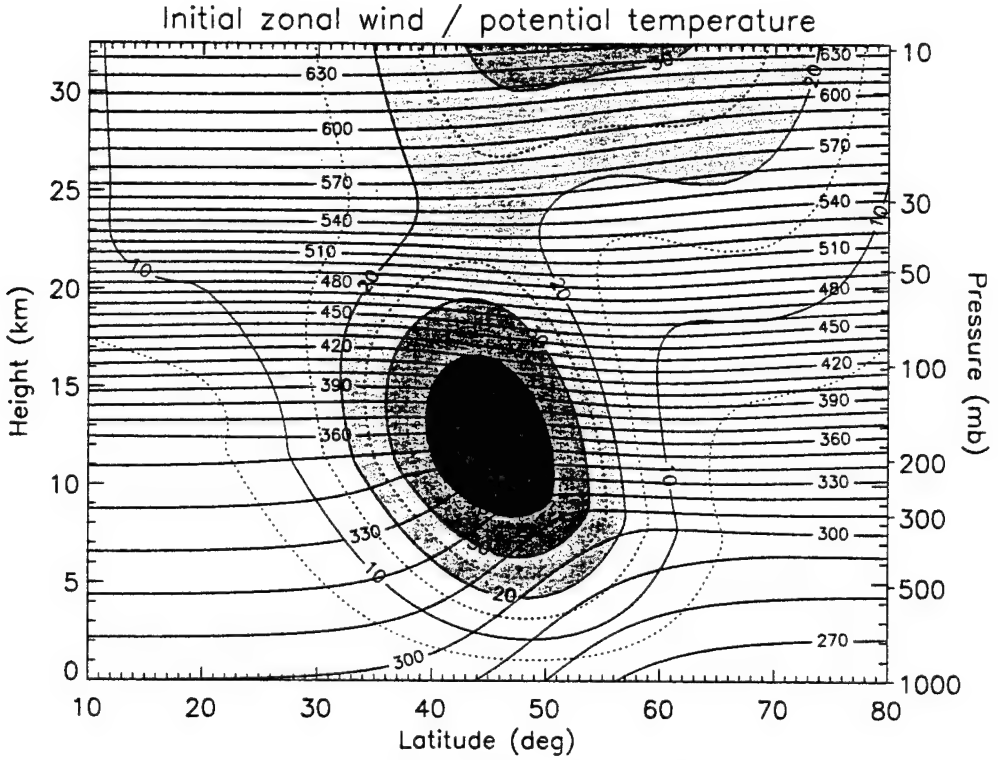


FIG. 1. Axisymmetric initial state used in model simulations, showing isopleths of potential temperature (solid contours ranging from 270 to 640 K) and mean zonal wind (shading and alternating solid/dotted contours).

and vertical resolution, for example, T42, T84, T126, and T156, and Δz = 350 or 700 m. Our custom has been to maintain approximately the same hyperdiffusion damping time for the smallest resolved scale, regardless of horizontal truncation. Higher horizontal resolutions therefore imply smaller hyperdiffusion coefficients and vice versa, unless stated otherwise.

Our primary interest was the generation of IGWs by geostrophic adjustment of the tropospheric jetstream as it is distorted by a developing baroclinic wave. For this purpose, the evolution should be realistic but not overly complicated. Following Simmons and Hoskins (1978), the model was initialized with a small-amplitude wavenumber 6 normal mode superposed on a zonally uniform, baroclinically unstable flow. This procedure yields within several days of model integration a well-resolved and fairly realistic finite perturbation to the jetstream, followed by a barotropic decay phase as described in section 3a. There are of course six identical lifecycles around the Northern Hemisphere.

For time integration the model used a semi-implicit scheme (Hoskins and Simmons 1975) with time step of 600 s, as required by the CFL condition for windspeeds attained in the simulation. As we were mainly interested in the inertia–gravity waves, and it is known that the semi-implicit scheme affects (high frequency) gravity wave phase speeds, two T84 simulations of the same lifecycle were compared using either the semi-implicit scheme or a fully explicit, centered-difference time integration scheme. These simulations used time steps of 1200 and 150 s, respectively (demonstrating the computational advantage of the semi-implicit scheme). Comparing simulations during the mature stage of the lifecycle showed identical lifecycles occurring. Development of the lifecycle's mature stage was delayed by about 1 day in the semi-implicit case, probably reflecting the different treatment of initial condition by the two schemes. Once this stage was reached, subsequent evolution did not differ noticeably. In particular, the low-frequency inertia–gravity waves of interest were insensitive to the choice of time integration scheme and appeared nearly identical in the two cases.

3. Baroclinic lifecycle and IGW excitation

a. Baroclinic lifecycle simulation

As described above, a baroclinic lifecycle analogous to that of Simmons and Hoskins (1978) was simulated with resolution adequate to model inertia–gravity waves. The unstable baroclinic wave developed on an initially axisymmetric jet centered at 45°N as shown in Fig. 1. The initial state contained several realistic features, namely, a change of static stability across the tropopause ($N^2 = 1.5 \times 10^{-4}$ s^{-2} in the troposphere,

$N^2 = 3.5 \times 10^{-4}\ s^{-2}$ in the stratosphere), a sloping tropopause, a stratospheric polar night jet, and windspeed minimum between stratospheric and tropospheric jetstreams. The initial state, with its strong jet displaced north of the climatological jetstream, might represent a cross section through midoceanic (Pacific or Atlantic) storm tracks during winter. It was disturbed by a small wavenumber 6 perturbation, which subsequently grew via baroclinic instability to finite amplitude. After 7 days of linear growth the baroclinic wave appeared as undulations on the westerly current.

Discussion of the baroclinic wave and accompanying IGWs begins on day 9 of the model run, by which time the baroclinic wave at upper levels attained finite amplitude. At midtropospheric levels an amplifying ridge, and deepening trough and low pressure center, show that the wave is still in its growth phase. Days 10 and 11 may be considered the wave's mature stage as growth ceases and the elongated SW-trailing trough sheds a cutoff low near 30°N (Fig. 2a). After formation of the cutoff low, the lifecycle enters its decay stage with the main westerly jetstream shifting northward on day 12 to 50°–60°N and becoming more zonally symmetric and intense (Fig. 2b). At this time the baroclinicity has been largely removed from midlatitudes where areas of weak horizontal temperature gradient predominate, flanked to the north and south by frontal zones.

At low levels the characteristic poleward heat flux of baroclinic instability is seen as the cold air, is advected southward and the warm air northward (Figs. 2c,d). Cyclogenesis and frontogenesis proceed rapidly with a surface low and warm/cold fronts well formed by day 9. By this time, the surface fronts have sharpened to the point that their width is limited by the model's small-scale hyperdiffusion.

The low-level cyclogenesis and frontogenesis strongly resemble the observed development of synoptic systems as discussed, for example, by Shapiro and Keyser (1990). The near-surface temperature on day 9 or 10 shows the so-called T-bone frontal stage of the baroclinic lifecycle, a sharp front at 60°–65°N extending back into the low pressure center, while the cold front, advancing eastward of the low, forms the vertical stem of the T. A pocket of warm air has been secluded behind the cold front near 55°N, 15°W on day 9 and 10, as has been observed over oceans when cyclones reach the mature stage of their lifecycle (apparently because of the lower surface friction there).

The baroclinic lifecycle also resembles that of Simmons and Hoskins (1980) and Thorncroft and Hoskins (1990), even showing the secondary frontal development discussed by the latter authors. This occurs on days 11 and 12 as the upper-level circulation associated with the cutoff low drifts over the surface cold front, initiating a small-scale frontal wave near 40°N, 20°W (Fig. 2d)—an example of Peterssen's "type B" cyclogenesis.

In the tropopause region the evolution is like that described for the midtroposphere. The jetstream is similarly distorted as the baroclinic wave radiates upward. After a period of linear growth, the wave saturates and a cutoff low forms by day 11. This is a period of rapid geostrophic adjustment throughout the troposphere but, especially, at upper levels where the jetstream is strongest.

We next examine evidence of gravity wave activity associated with surface frontogenesis and geostrophic adjustment in the upper troposphere. The field of horizontal velocity divergence is useful for isolating gravity waves from the large-scale quasigeostrophic flow because of the divergent nature of such waves. (By contrast, the signature of IGWs in geopotential is hardly visible.) The horizontal divergence field includes forced ageostrophic circulations accompanying the varying jetstream. As shown next, waves located away from this apparent source region show a radiating pattern of divergence distinct from that associated with the forced circulations.

b. IGW generation during lifecycle

The divergent component of tropospheric flow evolves quickly and in a complicated manner during the baroclinic lifecycle. Analysis of horizontal velocity divergence (hereafter $\nabla \cdot \mathbf{v}_H$) shows gravity wave activity associated with the surface fronts and cross-frontal ageostrophic circulations. These surface-related features weaken with increasing height because frontogenesis is strongest at ground level, and the radiating waves are dissipated as the mean flow turns with height. At the tropopause level and above (10–25 km), gravity wave activity increases dramatically during the lifecycle's mature stage. Figures 3a–d show $\nabla \cdot \mathbf{v}_H$ at 130 mb (near 14-km height) on days 9–12. The geopotential field (Figs. 4a–d) shows the location of the jetstream at this level. The first appearance of significant gravity wave activity in the upper troposphere occurs on day 9. Over the next day this activity quickly increases as a gravity wavepacket reaches 130 mb. The wavepacket is strongest on day 11, having advanced eastward with a secondary wave train splitting and propagating southward around the west side of the cutoff low. By day 12, the main packet has advanced rapidly eastward due to acceleration of the jet. A third wave train appears on the east side of the cutoff low, unrelated to the first packet. Over subsequent days, fragments of $\nabla \cdot \mathbf{v}_H$ become more widely spread and less coherent (not shown).

Figures 3 and 4 show that the upper-level gravity waves tend to be confined to the vicinity of the jetstream, propagating with wavevector parallel to the jet axis at midstream but at an angle to the jet along its flanks. This is also observed on the east and west sides of the cutoff low. Interestingly, at late times the remnants of IGWs in the cutoff low are carried back into the jet as the low itself is absorbed (not shown).

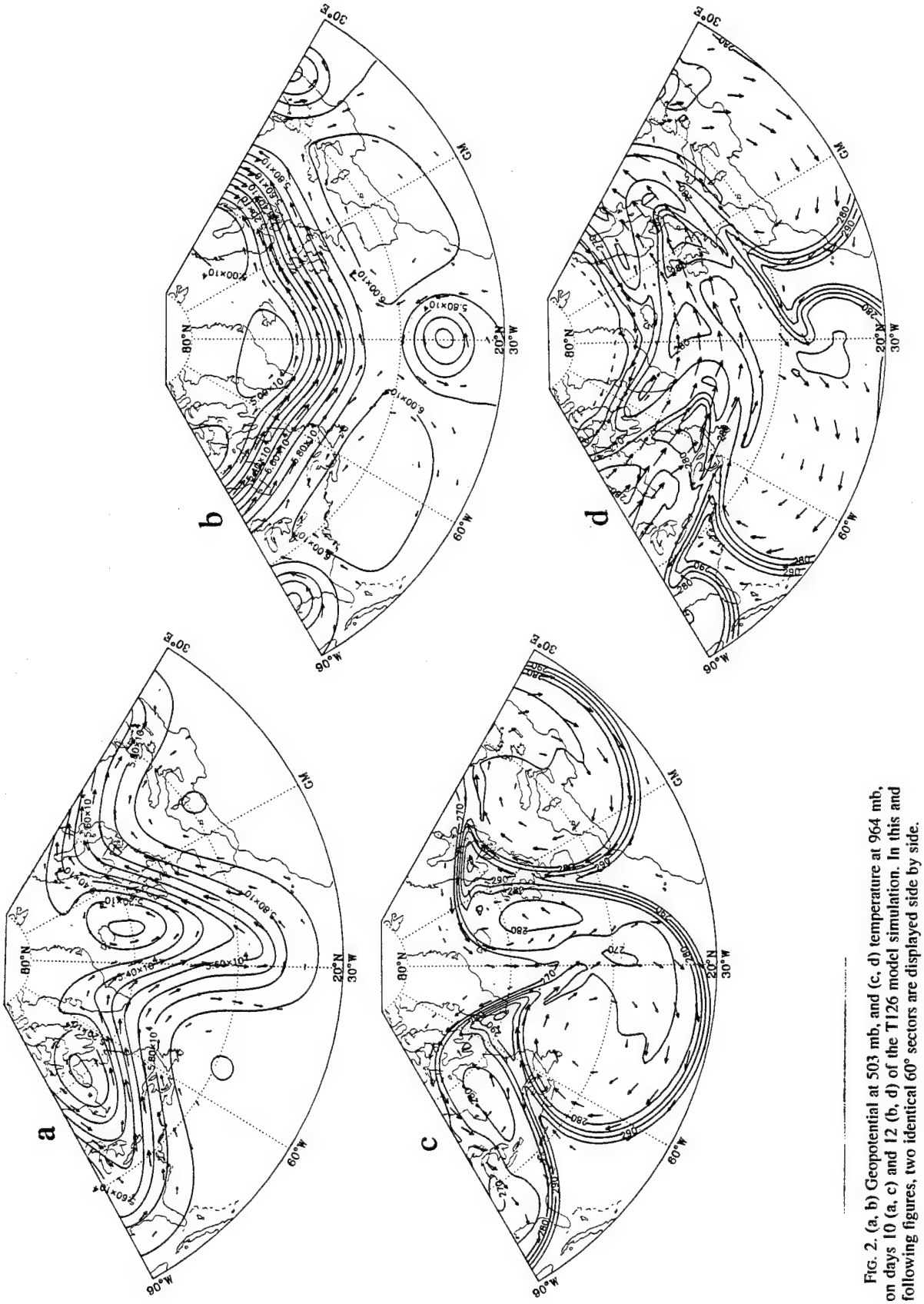


FIG. 2. (a, b) Geopotential at 503 mb, and (c, d) temperature at 964 mb, on days 10 (a, c) and 12 (b, d) of the T126 model simulation. In this and following figures, two identical 60° sectors are displayed side by side.

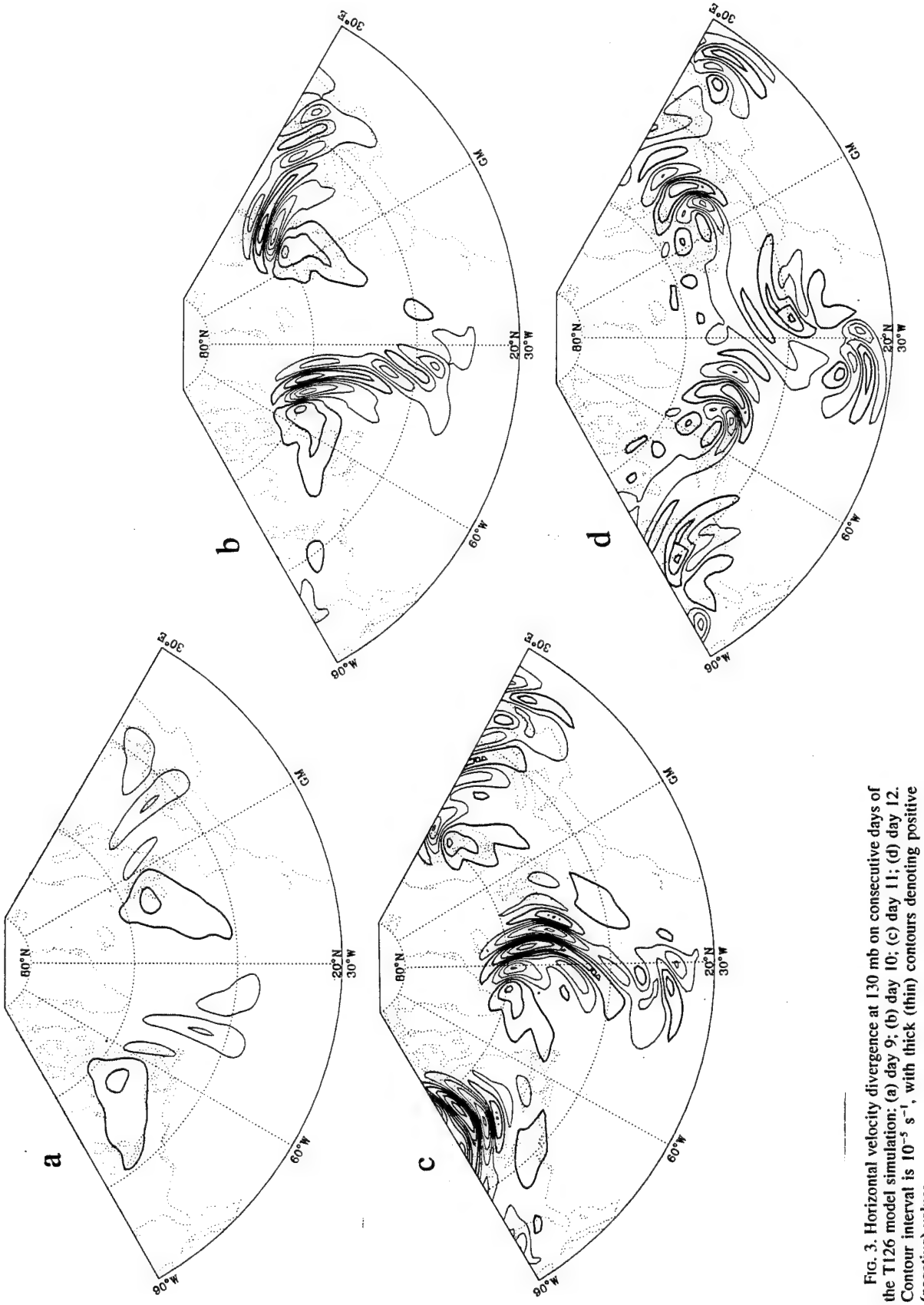


FIG. 3. Horizontal velocity divergence at 130 mb on consecutive days of the T126 model simulation: (a) day 9; (b) day 10; (c) day 11; (d) day 12. Contour interval is 10^{-5} s^{-1}, with thick (thin) contours denoting positive (negative) values.

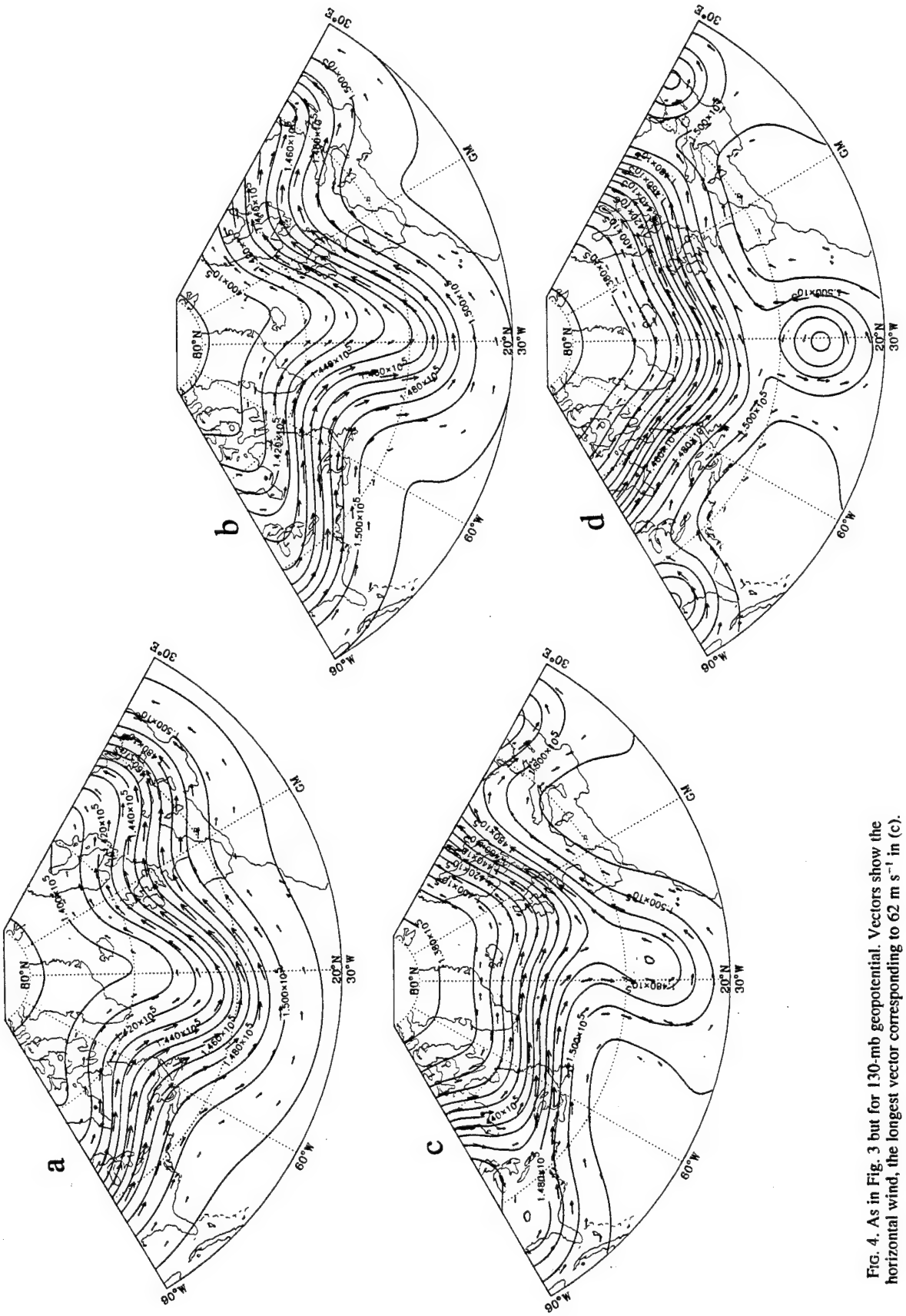


FIG. 4. As in Fig. 3 but for 130-mb geopotential. Vectors show the horizontal wind, the longest vector corresponding to 62 m s^{-1} in (c).

While model gravity waves in the mid- and lower troposphere can be related to frontal activity, those in the upper troposphere and lower stratospheric generally cannot and appear to have a different origin. In order to provide a 3D view of these waves, and to clarify their origin, cross-sectional views of the wavepacket are presented in Figs. 5 and 6. Figures 5a,b show latitude–height sections of $\nabla \cdot \mathbf{v}_H$ on days 10 and 11 at longitudes that intersect the gravity wavepacket on each day. In these cross sections the dominant feature is the enhanced gravity wave activity at upper levels, between the tropopause and ~20 km. Frontal features at low levels are also evident, largely unconnected to the activity aloft. The upper-level waves appear like a stack of pancakes, with phase fronts sloping down on either side. These surfaces also tend to parallel the isopleths of constant windspeed.

Figures 6a–d show longitude–height cross sections of $\nabla \cdot \mathbf{v}_H$ on days 9–12 at 55°N, a latitude intersecting the main wavepacket on those days. This again shows the upper-level wavepacket to be the dominant feature. The waves have phase surfaces sloping uniformly upward to the west. It is convenient here that the jetstream exit region,[1] located near the crest of the geopotential ridge, is propagating eastward almost parallel to a latitude circle during this stage of the lifecycle. Thus the longitude–height cross section at 55°N contains the exit region on these days. Figure 6 illustrates an important feature of the upper-level gravity waves, namely that their phase and group speed are small (or zero) relative to the jetstream exit region. This is best seen in animations of the cross-sectional view, but it can also be discerned from the plots for day 9–11, where the first three distinct wave crests (counting upward from the bottom of the wavepacket) are advected eastward with the jetstream exit region. In animation it appears that the jetstream pattern and attendant gravity waves are jointly advected eastward as a coherent structure. While the waves are approximately stationary relative to the jetstream pattern, they are, of course, in a strong westerly flow, so their intrinsic phase velocity is rapidly westward. The intrinsic group velocity is also westward as discussed below. (Likewise, the intrinsic propagation of the jetstream pattern is westward with respect to the fluid in the upper troposphere.) An upward group velocity can be discerned for these waves over the 3-day period of Fig. 6, as the wavepacket expands upward with time.

The impression gained from Fig. 6 is of gravity wave generation in the upper troposphere near the level of maximum wind in the vicinity of the jetstream exit region. There is a striking vertical asymmetry with most of the IGWs on the upper side of the jet, where static stability increases dramatically with height. Over the 3 days of the "mature lifecycle" stage the gravity waves radiate upward into the lower stratosphere where, except for one or two small packets, most of them approach a critical surface and are absorbed. (Note that critical layer absorption can occur when the mean flow turns with height—the windspeed need not approach the wave's phase speed.) There seems to be continual generation of gravity wave activity near the level of maximum wind during this time, since the lower edge of the wavepacket does not move upward. Van Tuyl and Young (1982) noted similar stationarity of unbalanced motions (IGW) relative to the jet pattern in their two-level mechanistic experiments of geostrophic adjustment. As noted by those authors, IGWs are a subsynoptic motion that can have significant amplitude in the exit region of strong jetstreams. In our high-resolution simulations IGWs explain a substantial part of the divergent wind component and at maximum amplitude (3–5 m s^{-1}) constitute about 10% of the total wind near the tropopause.

To test the conclusion that IGWs are due to geostrophic adjustment of the upper level jet and are not caused by gravity wave generation near the surface frontal regions, the simulation was repeated with strong damping of horizontal velocity divergence in the lowest 5 km. Applying *additional* sixth-order hyperdiffusion to divergence only—500 times stronger at the surface than the standard value and decreasing linearly to zero at 5 km—gave results similar to the standard case but with much smoother horizontal velocity divergence at low levels, effectively eliminating any gravity wave radiation from the surface fronts. At upper levels the flow was very similar except for a slight reduction of jet core speed (~5%). Most importantly, IGWs were again found, unchanged other than a slight reduction of amplitude (~10%).

The IGWs are well resolved in this T126 simulation, at least near the wavepacket center where their horizontal wavelength is 600–1000 km. This was borne out by comparison to an identical T84 simulation showing a very similar IGW packet with identical phase structure on day 11 at 130 mb. Differences were noticeable at the jet flanks where the IGW horizontal wavelength undergoes contraction as discussed in sections 3c,d.

c. *Description of IGWs at upper levels*

The gravity waves found in the simulated baroclinic lifecycle are of inertia–gravity type; that is, $N \gg |\hat{\omega}| \geq |f|$, where N is the Brunt–Väisälä frequency, $\hat{\omega}$ is the intrinsic wave frequency, and f is the Coriolis parameter. Examining the center of the main IGW packet between days 10 and 12 indicates that the IGWs have a period with respect to the ground of about 12–24 h and propagate eastward at a slower rate than the speed of the (westerly) jet. Following Dunkerton (1984), we

[1] The term "exit region" is used to describe the relatively weak, tightly curved flow downstream of the jet maximum. Parcels in this area do not decelerate to zero but are deflected southward into the west side of the developing cutoff low.

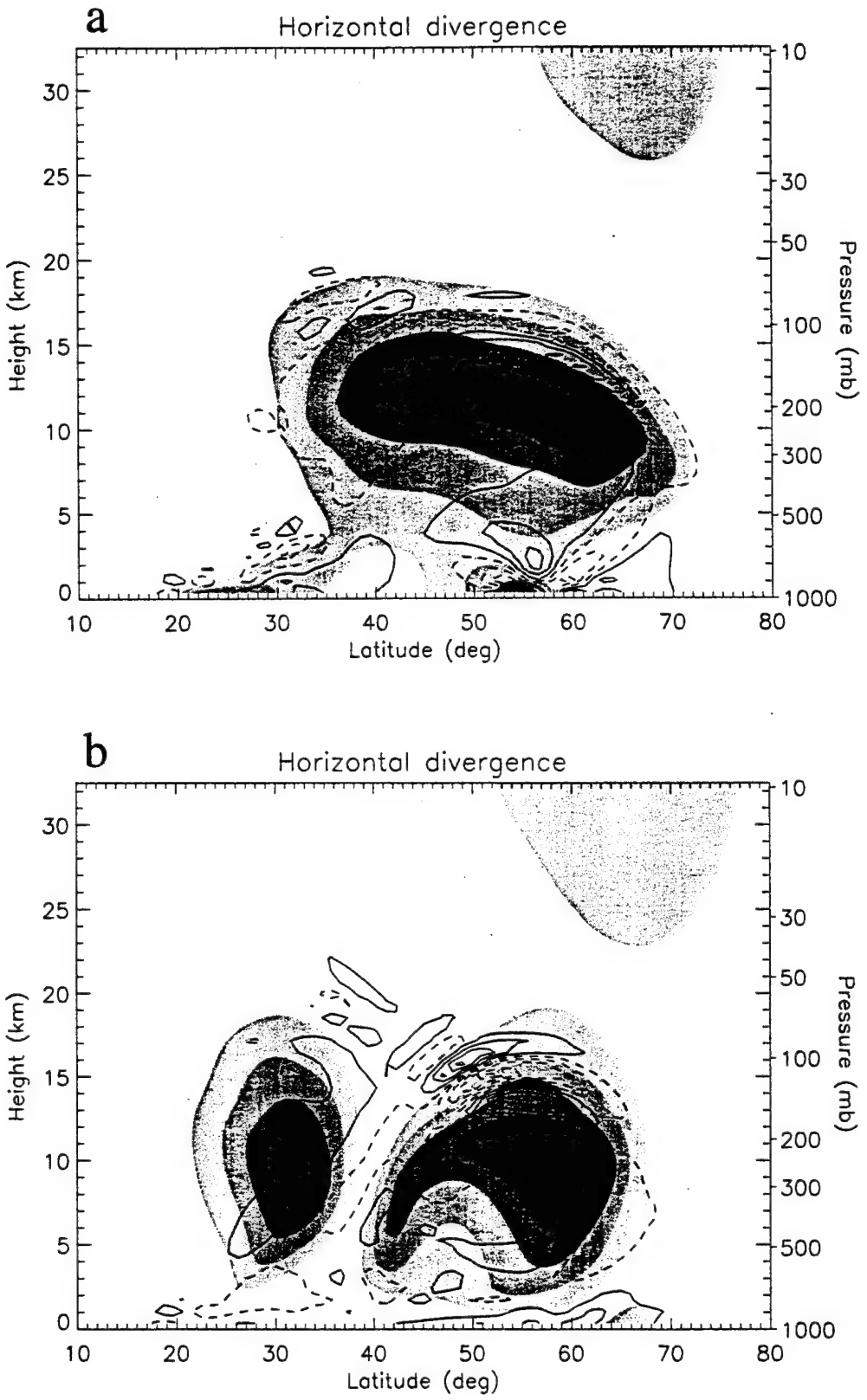


FIG. 5. Latitude–height cross section of horizontal velocity divergence at two longitudes following the main IGW packet on consecutive days: (a) day 10, 20°E; (b) day 11, 20°W. Contour interval is 10^{-5} s^{-1}. Windspeed superposed with shading increment 10 m s^{-1} as in Fig. 1.

/ ams v3628 1203 Mp 9 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203

adopt the convention that the wavevector points opposite the mean flow, reflecting the fact that phase propagation is westward with respect to parcels in the jet. Thus the intrinsic frequency is positive by convention, and the frequency relative to the ground is mostly in the range $-2\pi/\text{day} \leqslant \omega \leqslant -\pi/\text{day}$. For a zonally oriented wavevector with zonal wavenumber k and a zonal flow $\bar{u}$, the intrinsic frequency $\hat{\omega} = \omega - k\bar{u}$. Taking $k \approx -2\pi/(600\text{ km})$, $\omega \approx -2\pi/(14\text{ h})$, and $\bar{u} \approx 30\text{ m s}^{-1}$, representative of the zonal wind through the wavepacket at 130 mb on day 11, yields $\hat{\omega} \approx 1.9 \times 10^{-4}\text{ s}^{-1}$. Thus, $\hat{\omega} > f (\approx 1.2 \times 10^{-4}\text{ s}^{-1}$ at 55°N). The waves propagate westward relative to the wind with an intrinsic period of about 9 h, which is shorter than the inertial period of 14.6 h at this latitude.

This quick estimate of intrinsic frequency may be compared to that derived from the dispersion relation for plane hydrostatic IGWs of the form $\phi = \phi_0 \exp i(\mathbf{k}\cdot\mathbf{x} - \omega t)$; namely,

$$\hat{\omega}^2 = f^2 + \frac{N^2(k^2 + l^2)}{m^2 + 1/(4H^2)} \tag{3.1}$$

with l and m being the meridional and vertical wavenumbers, respectively, and H is the density scale height. (The second term in the denominator is negligible compared to the first.) On day 11 the wavepacket at 130 mb is centered near 55°N, and the horizontal mean wind and wavevector are both directed zonally. The horizontal wavelength is approximately 600 km with $k^2 \gg l^2$, and, from Fig. 6c, the vertical wavelength is about 4 km. Substitution into the dispersion relation yields $\hat{\omega} \approx 1.7 \times 10^{-4}\text{ s}^{-1}$, which agrees reasonably well with the other estimate. The two values of wave frequency at 130 mb indicate that gravity wave parameters are consistent with an IGW interpretation provided that estimates are made near the center of the wavepacket. Exact agreement is not expected, since the waves are not plane and the dispersion relation ignores latitudinal and vertical shear. A more detailed analysis of $\hat{\omega}$ is presented below.

The group velocity of the IGW packet may be derived from the dispersion relation:

$$\mathbf{c}_g = \frac{N^2}{m^2\hat{\omega}}\left[k, l, \frac{(k^2 + l^2)}{-m}\right] + \bar{\mathbf{u}}. \tag{3.2}$$

For wave parameters near the 130-mb jet axis (where $|k| \gg |l|$):

$$\mathbf{c}_g = (-7.1 \times 10^2, \sim 0, 4.7)\text{ km day}^{-1} + \bar{\mathbf{u}}. \tag{3.3}$$

The vertical group velocity is positive since m is negative. Upward group velocity is also indicated by a wind hodograph taken at day 11 through the center of the wavepacket at 53°N, 23°W (Fig. 7). The hodograph of perturbation winds (a five-point running mean having been removed) shows anticyclonic (clockwise) rotation of wind with increasing height, implying that the group velocity is directed upward. The major axis of the hodograph ellipse, which is parallel to the direction of wave propagation, is oriented from slightly south of east to north of west. The perturbation horizontal winds associated with the IGWs have magnitudes of about (5, 3) m s^{-1} for components parallel and perpendicular to the direction of wave propagation. At other locations we find similarly that the major axis of the hodograph ellipse is oriented in the direction of wave propagation suggested by Fig. 3 (not shown).

The horizontal group velocity is directed upstream relative to the wind and has a magnitude of about -8 m s^{-1}. The wavepacket is seen to propagate eastward at about 15–25 m s^{-1}, accelerating to the higher value after day 11. Although this agrees with predictions, the observed distribution of the wavepacket does not necessarily indicate the group velocity since it is affected by mechanisms of excitation and absorption, as well as propagation.

Upward radiating IGWs such as these will encounter a "Jones critical level" when $|\hat{\omega}| \to |f|$. Using wave parameters typical of the center of the packet, oriented in the direction of the mean flow, we estimate $\bar{u}_{CL} \sim 25\text{ m s}^{-1}$ as the critical mean wind speed. In this simulation IGWs are seen to penetrate slightly higher than the Jones critical level, to levels where the westerly windspeed has decreased into the range $\bar{u} \sim 10$–25 m s^{-1}. This penetration above the Jones critical level is likely due to transience in the forcing of the IGWs, the derived $\bar{u}_{CL}$ being appropriate for steady-state conditions. IGW breakdown is thought to occur via Kelvin–Helmholtz (KH) instability due to vertical shear of the wave's transverse wind component. Our large-scale model cannot resolve the KH instabilities, of course, so the IGWs are absorbed by hyperdiffusion as they approach the critical level. This could happen either through contraction of the vertical wavelength or rotation of the horizontal wavevector away from the direction of the mean flow, both of which would reduce the intrinsic frequency and vertical group velocity. We believe in this instance (as explained further in section 3d) that diffusion of model IGWs is due primarily to contraction of *horizontal* wavelength and rotation of the horizontal wavevector away from the direction of the mean flow. Comparing a T84 version of this simulation (with $\Delta z = 700$ m) to an identical T84 run with double vertical resolution ($\Delta z = 350$ m) gave strikingly similar IGW and vertical wavelengths. *Propagation of IGWs in the model is therefore not limited by vertical resolution.* Furthermore, by repeating the simulation at higher horizontal resolution while keeping the hyperdiffusion coefficient *unchanged* gave almost identical IGWs. Thus, the simulated IGWs are well resolved and are not significantly affected by inconsistent vertical or horizontal resolution as discussed by Lindzen and Fox-Rabinovitz (1989).

Along the flanks of the jet the horizontal wavevector is rotated inward in response to differential advection

by the jet. As seen in Fig. 3, the horizontal wavelength along the flanks of the jetstream has contracted somewhat relative to that along the jet axis. This is consistent with waves refracting into the jetstream as discussed by Dunkerton (1984); for example,

$$\frac{dl}{dt} \approx -k\bar{u}_y, \tag{3.4}$$

etc. On the southern flank of the jet $l > 0$ so $c_{g_y} > 0$, while $l < 0$ and $c_{g_y} < 0$ on the northern flank.

A consequence of rotation and contraction is that wave activity along the jet flanks is rapidly dissipated by the model's hyperdiffusion. This mainly accounts for the sensitivity of simulated IGWs to the model's horizontal resolution: at lower resolution (T84), IGW activity away from the jet axis is lost even more readily to diffusion. We expect that with finer horizontal resolution well beyond T126, the trailing IGWs along the flanks of the jet would be resolved for a longer time, evolving to even shorter horizontal wavelengths. Some evidence in support of this claim is provided by results of a T156 run as discussed in section 3d.

That the waves are being absorbed along the lateral boundaries of the wavepacket is also suggested by calculations of intrinsic frequency, determined objectively from the spatial distribution and temporal behavior of $\nabla \cdot \mathbf{v}_H$. Wavevector orientations and magnitudes were calculated at approximately 1° intervals within the wavepacket (at 130 mb) by applying a high-pass spatial filter to $\nabla \cdot \mathbf{v}_H$ (to eliminate large-scale ageostrophic circulations) and evaluating the horizontal gradient of this quantity at the zero crossing. Although the phase function is ill behaved in a complicated wave field such as this, a local wavevector direction may be determined uniquely as parallel (or antiparallel) to $\nabla(\nabla \cdot \mathbf{v}_H)$ at the zero crossing. Half-wavelengths were estimated as the distance between adjacent zero crossings in the direction of the wavevector. To get reasonably smooth results, individual estimates within a 4×4 matrix of grid points were binned together and interpolated back to the original grid. A similar binning procedure was used for frequency ω, crudely estimating the local half-wave period as the time interval between adjacent zero crossings in a bandpass-filtered time series of $\nabla \cdot \mathbf{v}_H$ at each grid point. The wavevector direction and sign of ω were carefully determined by examining the phase propagation around the time of interest. Wavevectors were aligned in the direction of intrinsic phase propagation, that is, with a component antiparallel to the mean flow. The sign of ω was negative in all but a few places; this quantity was typically smaller than the advective term $\mathbf{k}\cdot\bar{\mathbf{u}}$ by a factor of 3–4. To evaluate the intrinsic frequency $\hat{\omega} = \omega - \mathbf{k}\cdot\mathbf{u}$, a small amount of spatial smoothing was applied to the horizontal wind, since IGWs are not part of the "mean flow."

Figures 8a–c show $\nabla \cdot \mathbf{v}_H$ (with high-pass spatial filter) at 130 mb on days 9.5, 11, and 12.5, illustrating the early, middle, and late stages of IGW evolution at this level, respectively. In the early stage (Fig. 8a) there is a single wavepacket in the jet stream exit region, with most wavevectors pointing backward and to the right with respect to the local mean flow (for an observer facing downstream). By day 11 (Fig. 8b), the packet has begun to split, part of it going south into the cutoff low. Wavevectors in both halves converge into the jet, and a few waves manage to sneak across the saddle of the cutoff low into a region they do not belong. By day 12.5 (Fig. 8c), the main part has advected rapidly eastward to a point immediately northwest of the cutoff low (reentering from the left side of the diagram), while the wavepacket trapped within the cutoff low remains largely stationary. Waves crossing the saddle rapidly disappear, replaced by a third packet (unrelated to the others), on the east side of the cutoff low, on a collision course with the first.

The intrinsic frequency on day 9.5 (Fig. 8d) shows most values lying in the range f–$4f$, with a flat ledge of smaller values (slightly below f) along the northeastern boundary of the wavepacket. The variation of $\hat{\omega}$ is due mainly to the angle of wavevector with respect to the mean flow and not the wavelength (which is rather uniform on this day). Although penetration beyond $\hat{\omega} = f$ would require that a steady-state wave be evanescent, in this case the phenomenon does not persist (as noted above, in connection with the vertical propagation) and is probably due to transients excited by geostrophic adjustment. On day 11 (Fig. 8e) the region of subinertial frequency has disappeared, although by this time there is another such region lying within the saddle of the cutoff low, which is likewise temporary and quickly damped. Most values of $\hat{\omega}$ are again in the range f–$3f$. By day 12.5, a larger percentage of wavevectors are becoming nearly orthogonal to the mean flow, and values of $\hat{\omega}$ are somewhat smaller, mostly f–$2f$ (Fig. 8f).

It must be noted that the absorbing action of the sidewalls of the jet is an artifact of the model, being sensitive to the model's horizontal resolution and hyperdiffusion. These results cannot address the (still unanswered) issue of whether the surface $\hat{\omega} = f$ acts as a critical layer to absorb IGWs (as in vertical shear) or as a turning point to reflect them (as in tidal theory). This matter was discussed by Kitchen and McIntyre (1980) among others but deserves further study. There is little, if any, evidence for reflection of model IGWs, although such an effect might be masked by the rapid evolution of the flow. It was suggested by Dunkerton (1984) that, due to horizontal and temporal variations of the basic state, any natural waveguides for IGWs are likely to "leak" wave activity rather than contain it. In the simulations some wave activity manages to escape to the stratosphere, but scale-dependent damping is the dominant effect.

Radiation to the middle stratosphere occurs only in the trailing part of the main wavepacket, where IGWs

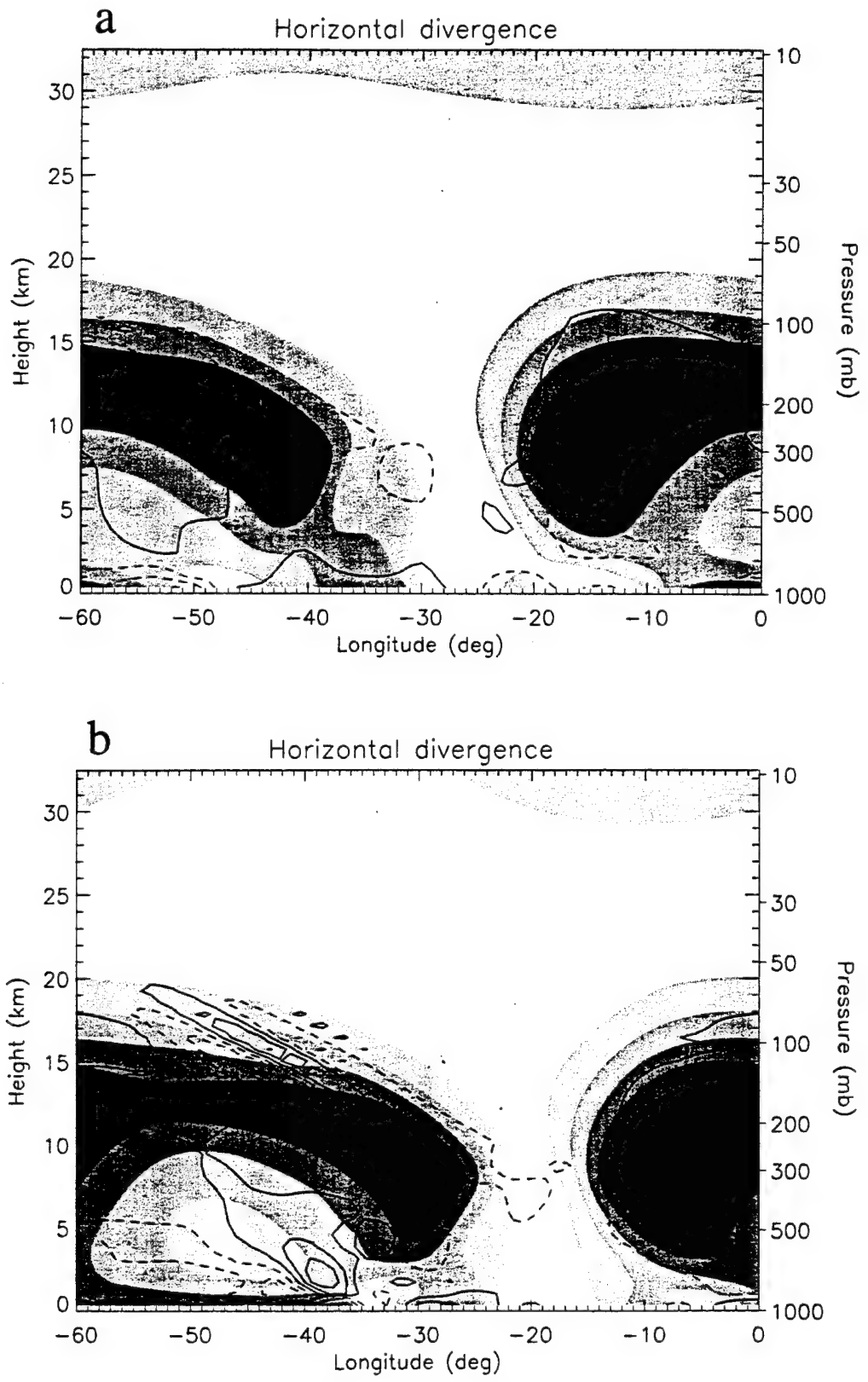
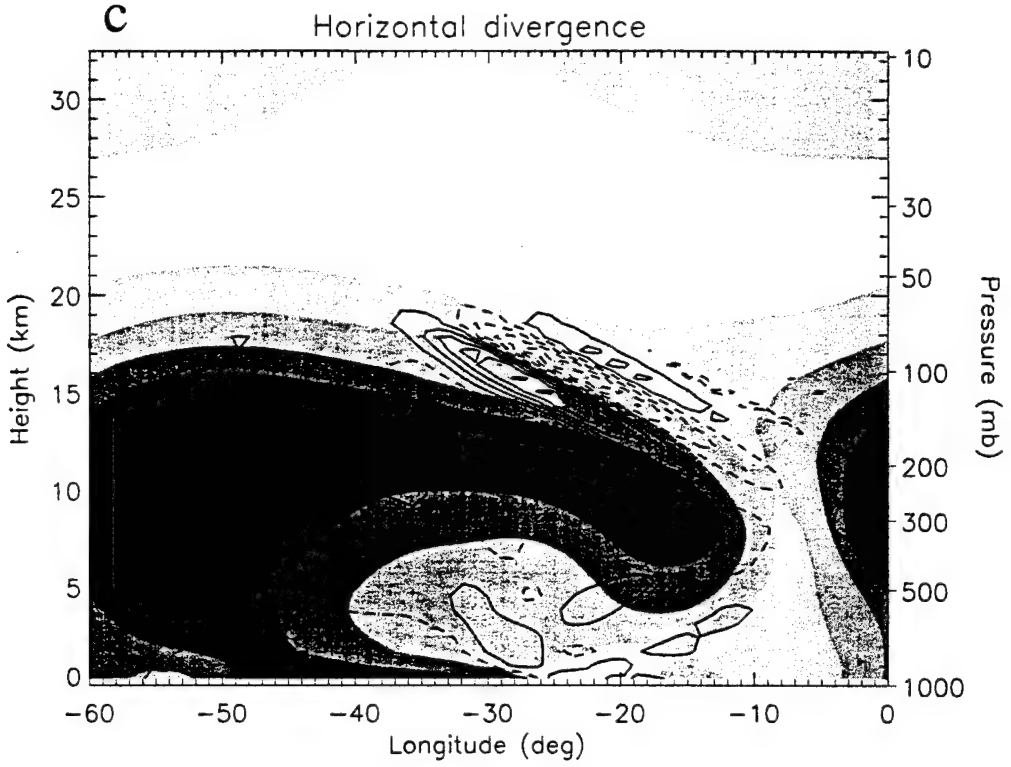


FIG. 6. Longitude–height cross section of horizontal velocity divergence at 55° latitude on consecutive days: (a) day 9; (b) day 10; (c) day 11; (d) day 12. Contour interval is 10^{-5} s^{-1}. Windspeed superposed as in Fig. 5 with contour interval 10 m s^{-1}.

/ams v3628 1203 Mp 12 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203



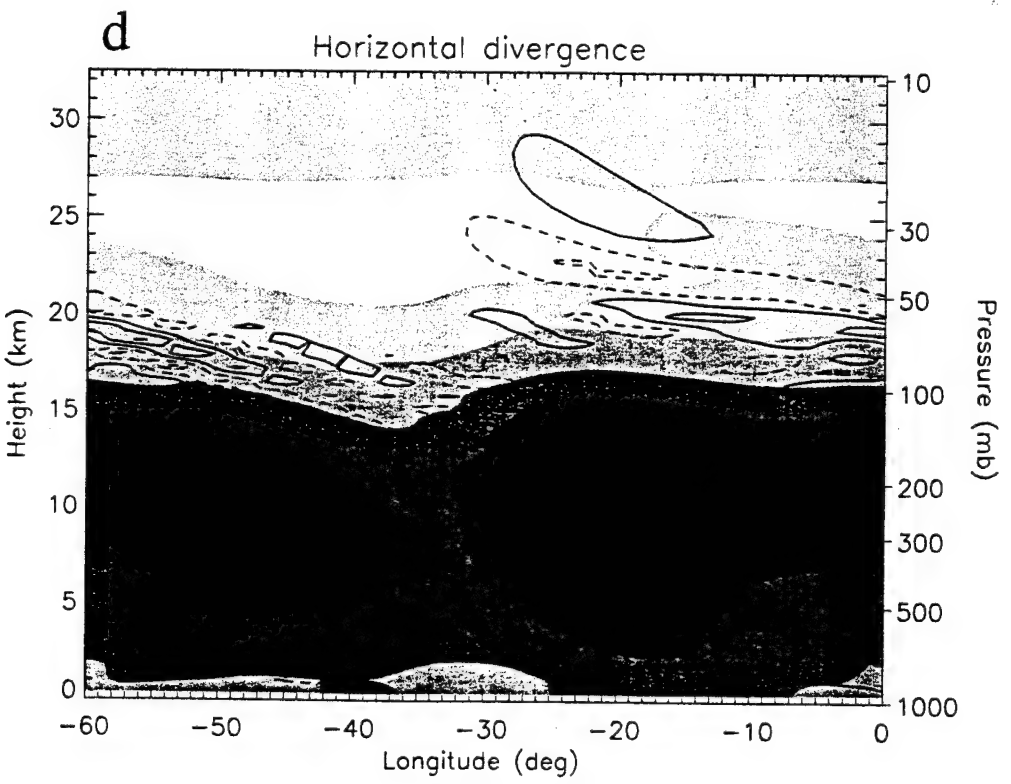


FIG. 6. (*Continued*)

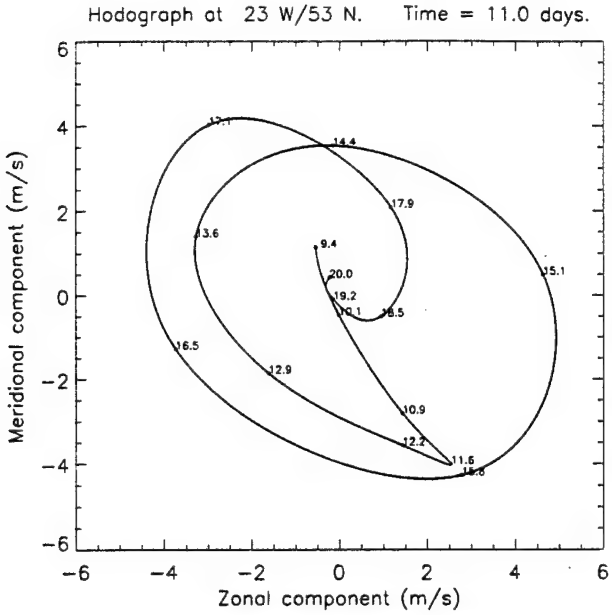


FIG. 7. Hodograph of the vertical profile of perturbation horizontal wind components on day 11 at 53°N, 23°W. Perturbation denotes deviation from winds smoothed with a five-point running mean vertically. Altitude of model levels indicated in km.

tilt northward into the polar night jet (not shown). Propagation at this location is possible on account of increasing mean windspeed despite the wavevector becoming nearly orthogonal to the mean flow. Another packet approaches 20 km on the west side of the cutoff low (not shown). These results demonstrate that vertical propagation of model IGWs is limited to the region of strong mean windspeed where tropospheric and stratospheric flows have approximately the same direction. The baroclinic wave distorts the tropospheric flow from a zonal orientation but is itself evanescent in height (i.e., does not similarly distort the middle stratosphere). Only a few tropospheric IGWs can therefore propagate into the midstratospheric circumpolar vortex.

d. Effects of horizontal and vertical resolution

Model IGWs are unaffected by vertical resolution if the grid spacing is sufficiently fine to describe the rotation of perturbation velocity with height, as shown in Fig. 7. Smaller Δz did not affect the evolution of model IGWs, indicating that the IGWs are primarily sensitive to horizontal hyperdiffusion.

Integrations were performed at horizontal resolution T42, T84, T126, and T156 with hyperdiffusivity coefficients $\nu_6 = 1.37 \times 10^{27}$, 1.02×10^{26}, 1.36×10^{25}, 2.52×10^{24} m^6 s^{-1}, respectively, such that the damping time at the smallest resolved scale was nearly the same in each run (~20–30 min). These integrations illustrate quite dramatically the sensitivity of model IGWs, as shown in Figs. 9a–d. Baroclinic lifecycles in each run were at virtually the same stage at day 10, but major differences are seen in model IGWs. Contour intervals are the same in each panel, indicating a large increase of IGW amplitude between T42 (when IGWs were largely absent) and T126. Figure 9 also demonstrates that certain parts of the divergence pattern are adequately resolved at lower resolution, while others are not. Increasing horizontal resolution beyond T84 did not significantly affect the horizontal scale of IGWs at the rear of the wavepacket nor the structure of forced ageostrophic circulations evident at T42. IGW structure at the center of the wavepacket was largely unchanged going from T126 to T156. Differences were observed mainly on the fringes of IGW packets, for reasons discussed in the previous subsection.

Interpretation of these results is complicated by possible changes of the baroclinic wave as resolution is increased, for example, sharper curvature of the jet and stronger gradients of velocity. An example of windspeed at day 10 for T84 is shown in Fig. 9e. Windspeeds and velocity gradients were slightly stronger at higher resolution, by a few percent. The basic pattern of horizontal velocity divergence within the source region (Fig. 9f) was similar, except for substantially more finestructure at higher resolution due to IGWs (not shown). Therefore, changes of the jetstream pattern were minor compared to the observed changes of IGWs, leading us to believe that the primary effect of increasing resolution in this range (T42–T156) is to allow the region of imbalance to project better onto the IGW manifold, irrespective of any change in IGW sources. We cannot rule out the possibility that such sources may contract further in horizontal scale or become more intense locally, as resolution is increased beyond T156, with a resulting contraction of IGWs and further increase of amplitude.[2] Indeed, upper-level frontogenesis is expected to excite mesoscale gravity waves. Model experience suggests, however, that such waves would add to the spectrum of IGWs already resolved. Important to this argument is the assumption that model IGWs are attributable to an adequately resolved region of imbalance and do not arise entirely from frontal collapse or from instabilities lacking a preferred horizontal scale.

e. Comparison with observations

The upper-level IGWs seen in this model study are broadly consistent with observations of IGWs near midlatitude jetstreams, except possibly regarding their direction of horizontal energy propagation. Typically, the lower stratospheric gravity wave field is dominated by quasi-monochromatic waves with horizontal wave-

[2] Requirements for vertical resolution become more stringent as horizontal hyperdiffusion is decreased, allowing closer approach of IGWs to critical levels.

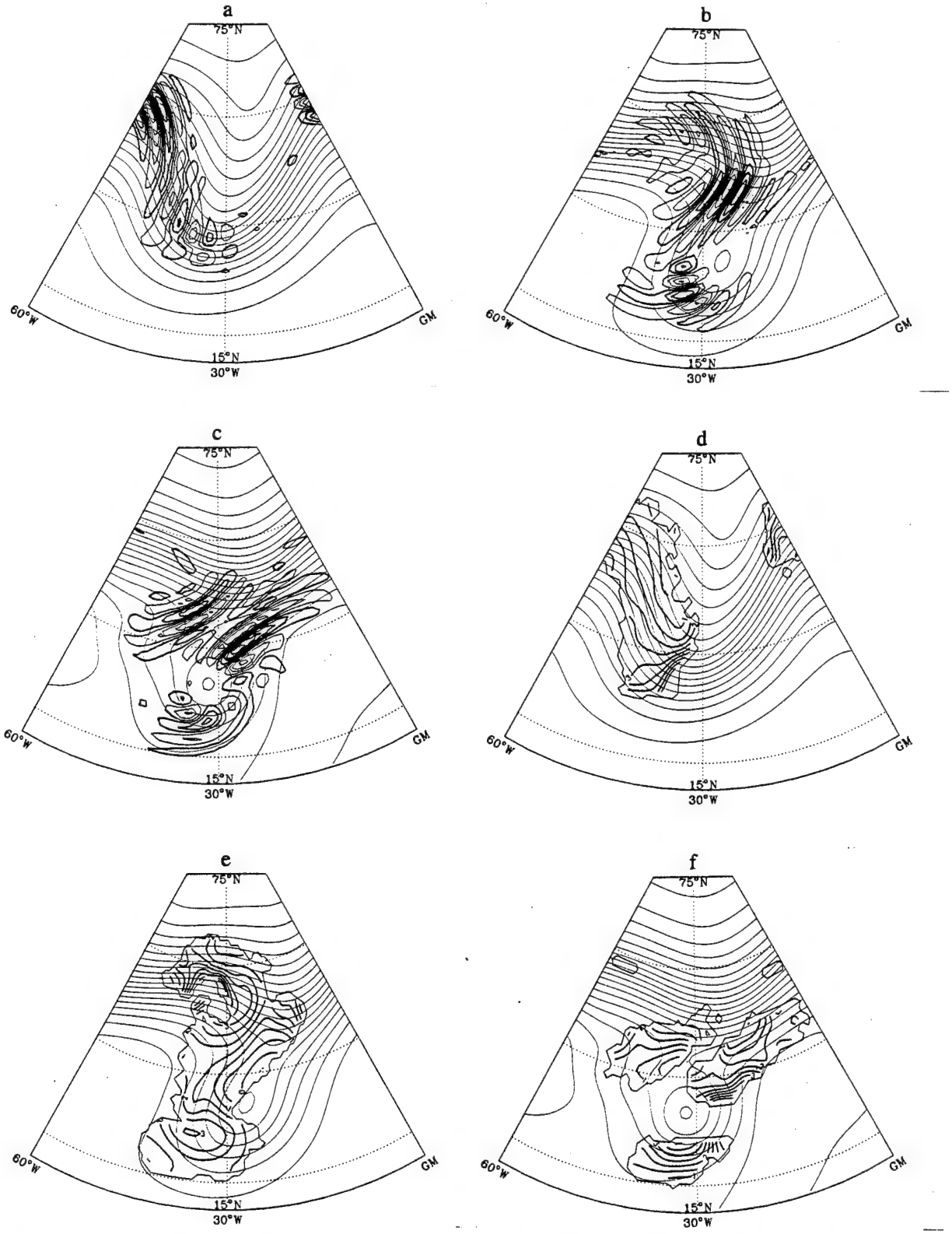


FIG. 8. (a–c) Pattern of horizontal velocity divergence (with high-pass spatial filter, each plot normalized by its maximum amplitude) and (d–f) normalized intrinsic frequency $\hat{\omega}/f$ of IGWs at 130 mb on day 9.5 (a, d), day 11 (b, e), and day 12.5 (c, f). Isopleths of geopotential superposed as thin contours, indicating the location of jetstream and cutoff low. Contours of filtered, normalized divergence should not be confused with wave amplitude, which is much smaller at early times.

/ ams v3628 1203 Mp 15 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203

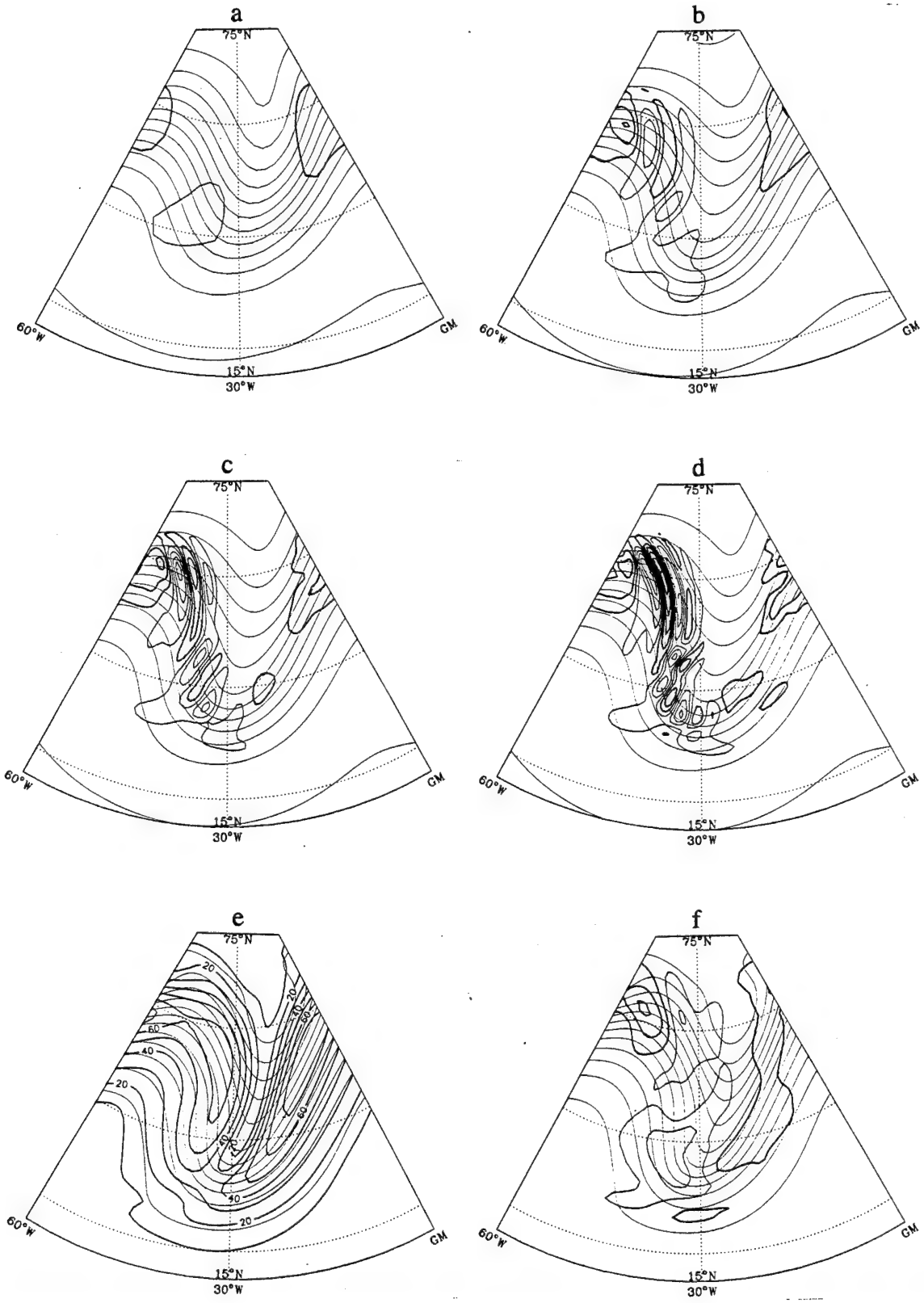


FIG. 9. Horizontal velocity divergence at day 10, 130 mb from various integrations: (a) T42; (b) T84; (c) T126; (d) T156. Contour interval as in Fig. 3. Thin contours display geopotential height at this level. (e, f) Windspeed and horizontal velocity divergence at day 10, 238 mb, with geopotential at this level superposed, from T84 run.

/ams v3628 1203 Mp 16 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203

lengths of several hundred kilometers and vertical wavelengths of 1–4 km with upward energy propagation. These IGWs appear to be related to the tropospheric jet stream, though it has proven difficult to directly relate them to a specific source mechanism. Thomas et al. (1992) reported a case study based on VHF radar observations made at Aberystwyth, Wales (52°N, 4°W). They found southward IGW energy propagation away from the jet stream, which was passing to the north. Prichard and Thomas (1993) examined stratospheric IGWs over 23 days and also noted the dominant southward IGW propagation direction. In apparent contrast to this are the results of Sato (1994), who statistically analyzed IGWs observed by radar at Shigaraki, Japan (35°N, 136°E). She found the lower stratospheric IGWs to be most energetic in wintertime and to propagate southward toward the strong subtropical jet, which is persistently located just south of there during winter. Finally, in situ observations of IGWs made with an ER-2 aircraft off the California coast in April 1986 found large-amplitude IGWs in a jet stream exit region with propagation roughly parallel to the jet axis (X. Pfister 19XX, personal communication). At 13–15-km height the vertical wavelength of these waves was 1–2 km, and the horizontal wavelength was deduced to be about 250 km.

Thus, while lower stratospheric IGWs primarily tend to propagate meridionally, there is some uncertainty about their relation to the jetstream. There may be large geographical variability in the relation of IGWs to nearby jetstreams. Similarly, it is likely that there are several different types of jet stream developments that result in significant IGW generation. For zonally elongated imbalances to the jet stream, meridionally radiating (away from the jet) IGWs should result. However, when the zonal extent of the imbalance region is reduced relative to its meridional extent, then zonally propagating IGWs should become significant (Luo and Fritts 1993).

Although not discussed in section 3c, the simulation at a later time shows evidence of zonally aligned IGW phase fronts propagating sideways into the tropical lower stratosphere. The waves are quite weak compared to those near the jet, but their (N–S) wavevector orientation is consistent with IGW observations at Arecibo Puerto Rico, described by Sato and Woodman (1982) and later authors. This would suggest that not all observed IGWs in the Tropics are due to underlying convection but that some of them may originate in the midlatitude jet.

f. Generation of IGWs in the upper troposphere

Given that IGWs are difficult to observe in atmospheric data lacking fine vertical and temporal resolution, it would be desirable to relate the excitation of IGWs to observable parameters in the large-scale flow field. The following thoughts are based partly on earlier literature and partly on our modeling experience, suggesting how this might be done.

As discussed in section 1, IGW generation is expected during periods of geostrophic adjustment of the tropospheric jet stream accompanying synoptic-scale developments. As the flow evolves, imbalances develop between the mass, temperature, and velocity fields causing ageostrophic motions that cannot be supported in a balanced system (quasi-geostrophic, semigeostrophic, nonlinear balance, etc.). In this event the balance breaks down, and IGWs are radiated away from the region of imbalance. There is no completely reliable way to diagnose the conditions required for IGW generation however, largely for lack of knowledge of the balanced state to which the flow is relaxing (Koch and Dorian 1988).

The horizontal wind divergence field near the level of maximum windspeed (~250 mb) shows the signature of slowly varying ageostrophic circulations that accompany the jet stream (Fig. 9f). These patterns are similar to those analyzed by Cammas and Raymond (1989) for jet streams where flow curvature was strong, for example, their Fig. 6. In such cases the familiar quadrupole pattern of upper-level divergence associated with the entrance and exit regions of a localized zonal jet stream are obscured by the divergence associated with the flow curvature. The appearance of a strong divergence center in the jet exit region at 250 mb, just upstream of the ridge, implies that air parcels rapidly advected through this region will experience a large local rate of change of divergence, suggestive of flow imbalance. In analogy to wave generation by flow over an obstacle, it is to be expected that the dominant waves generated by flow through this jet exit region should propagate with a wavevector approximately parallel to the jet axis, as observed in the model near the center of the jet. Also consistent with this idea, we find that the IGW pattern is approximately phase locked to the jet exit region over a period of several days (Fig. 6). Similarly, the IGW structure in the jet exit region is fairly insensitive to improvements in resolution.

Typical measures of imbalance rely on estimating the magnitude of those terms that are assumed to be negligible in any particular balanced-flow state. Thus, one might have a small Rossby number ($R_o = U/fL \ll 1$) for geostrophic balance, a small Lagrangian Rossby number ($R_o^{(L)} = |d\mathbf{V}/dt|/f|\mathbf{V}| \ll 1$) for semigeostrophic balance (since it implies $d\mathbf{V}/dt \sim d\mathbf{V}_g/dt$), and a small material derivative of horizontal divergence ($d\nabla\cdot\mathbf{v}_H/dt \ll 10^{-9}\,\mathrm{s}^{-2}$) for the nonlinear balance equation.

In this study the plots of $d\nabla\cdot\mathbf{v}_H/dt$ tended to mimic those of $\nabla\cdot\mathbf{v}_H$ (e.g., Fig. 3), which we use to show the IGWs. This quantity is therefore an accurate but redundant estimate of imbalance. Instead, one would like a measure of flow imbalance that does not explicitly involve $\nabla\cdot\mathbf{v}_H$ and may be diagnosed from the observ-

able flow. The Lagrangian Rossby number appears suitable for this purpose.

When applied to synoptic data, Koch and Dorian suggest an indicator of unbalanced flow is the Lagrangian Rossby number $R_o^{(L)} \geq 0.5$ in a region where $|\mathbf{V}| > 10$ m s^{-1} (to avoid spuriously large values of $R_o^{(L)}$ from the denominator being small). In practice, when calculating $d\mathbf{V}/dt$ from synoptic data, the local wind tendency $\partial\mathbf{V}/\partial t$ must often be neglected relative to the advective tendency $\mathbf{V}\cdot\nabla\mathbf{V}$ because of the crude time resolution of the observations. For this simulation we calculated both terms of $d\mathbf{V}/dt$ and found that, indeed, the local wind tendency is usually small compared with the advective tendency $\mathbf{V}\cdot\nabla\mathbf{V}$, although locally it can amount to ~30% of the latter.

Figure 10 shows $R_o^{(L)}$ at 195 mb on day 11 of the baroclinic lifecycle with horizontal velocity divergence overplotted at 130 mb, showing the main IGW packet at 50°–60°N and a second wavepacket west of the cutoff low near 30°N. Nearer to the jet core, $R_o^{(L)}$ was calculated at a slightly lower level, since the wave source lies below where the waves are prominently seen. In this simulation large values of $R_o^{(L)}$ are found extending down to the level of the jet maximum, near 250 mb. In Fig. 10, these large values (heavy lines) are seen to coincide with the location of IGW activity. Two regions of large $R_o^{(L)}$ occur: one between 45° and 60°N that overlies the main IGW packet in the jetstream exit region and a second region lying southwest of the cutoff low, suggesting that two separate regions of imbalance exist. To be sure, large $R_o^{(L)}$ suggests where an imbalance exists but says nothing about the intensity of resulting IGWs. Since IGWs radiate unbalanced energy away from the jet, their intensity should be approximately proportional to $|\mathbf{V}|^2$. This is consistent with results of mechanistic numerical experiments by Van Tuyl and Young (1982), who found that the strength of unbalanced motions (IGWs) increases roughly as the square of jet core strength. Thus, stronger IGW radiation is to be expected from regions of imbalance (large $R_o^{(L)}$) where windspeeds are strong. This may explain why the IGWs are much stronger in the high-latitude region of imbalance than at low latitudes where windspeeds are weaker.

A combination of Lagrangian Rossby number and mean flow kinetic energy, then, provides a useful diagnostic of IGW excitation in the model. It remains to be seen whether it is systematically useful in observations.

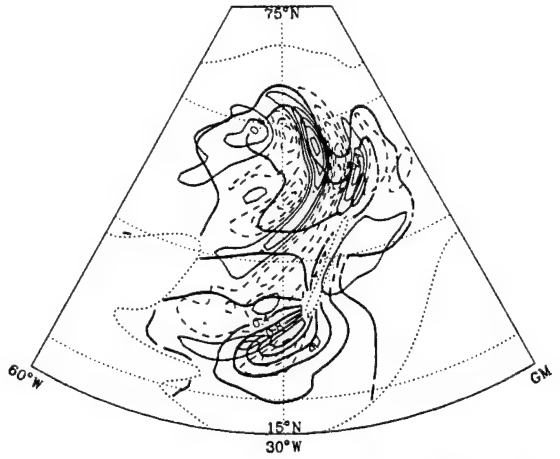


FIG. 10. Lagrangian Rossby number $R_o^{(L)}$ at 195 mb on day 11 (heavy lines), with horizontal velocity divergence at 130 mb superposed: $R_o^{(L)}$ is plotted only where the windspeed exceeds 10 m s^{-1}, the dotted line denoting this boundary.

4. Discussion

The ubiquitous nature of IGWs in the atmosphere has long been recognized, yet their effects are only beginning to be considered in the context of the general circulation. This situation is partly due to their subsynoptic scale and preferred occurrence at upper tropospheric levels and above. IGWs are nevertheless likely to play an important role in several aspects of atmospheric circulation.

Our simulations show that IGWs are generated in the upper troposphere, propagate energy upward, and are absorbed approaching the Jones critical level (either in weak winds or as the wavevector rotates away from the mean flow). As discussed in section 3c, IGW breakdown is thought to occur via KH instabilities (unresolved in the model), which develop as amplitude increases and vertical wavelength decreases. These instabilities are expected in the region of strong transverse vertical shear, at which point the flow is statically stable (Dunkerton 1984; Fritts and Rastogi 1985). IGW breaking is, therefore, a potentially efficient means of cross-isentropic transport and may be responsible for most of the vertical diffusion found in the lower stratosphere, a region where systematic cross-isentropic transport due to diabatic heating is quite weak. Vertical mixing due to IGW breaking must vary geographically and seasonally, just as IGW generation does.

In addition to cross-isentropic transport induced by IGWs, there is an isentropic component due to advection by the horizontal velocity. One normally thinks of gravity wave motions as fast, reversible undulations superposed on the "slow manifold" of rotational motions (e.g., nonlinear baroclinic instabilities and breaking planetary Rossby waves), the latter being mainly responsible for isentropic mixing on a global scale. Parcel displacements due to IGWs, nevertheless, can in certain cases be large, due to their large horizontal scale and low intrinsic frequency—particularly as the critical level is approached and amplitude increases. In this case an air parcel advected through the IGWs will undergo a considerable horizontal excursion during a half wave period before the transverse velocity reverses sign.

Under such circumstances, any background tracer field with a nonzero gradient orthogonal to the IGW wavevector will experience considerable deformation. This is true, in particular, for potential vorticity (PV) near the tropopause. Such deformations of PV on an isentropic surface are in fact seen in these simulations. Figure 11 shows PV on the 430-K surface, illustrating large meanders of the (approximately conserved) PV contours across the jet in a region of strong IGWs. The meanders are rapidly strained to smaller scales by the background flow and lost because of limited model resolution. It is certain that they will lead to irreversible isentropic transport and that this will be especially strong in the cross-jet direction if real IGWs behave like those simulated here. In ongoing work we are investigating the effect of IGWs on isentropic tracer transport in the lower stratosphere by comparing transport and mixing in simulations with and without IGWs. This is relevant in light of recent contour advection studies (Plumb et al. 1994; Waugh et al. 1994; Waugh and Plumb 1994; Norton 1994) in which the isentropic tracer redistribution is computed from advecting winds known only at low resolution and as infrequently as daily. Recognizing the importance of unbalanced motions, Pierce and Fairlie (1993) speculate on the role of IGWs in enhancing cross-jet tracer dispersion, perhaps enhancing the permeability of the lower stratospheric vortex at its edge.

Figure 12 shows a latitude–height cross section through the IGW-induced meanders shown in Fig. 11. This illustrates the effect of quasi-horizontal isentropic IGW motion on the PV field, which is representative of a conserved, vertically stratified tracer. An important implication of such isentropic transport in general is that IGWs may directly cause stratosphere–troposphere exchange. If the IGW breaking level occurred a few kilometers lower than seen in Fig. 12, the quasi-horizontal motions would involve the tropopause [commonly defined by the 2-PVU (potential vorticity units) contour]. For instance, if the IGW motions of Fig. 11 were to occur 100 K lower at approximately 330 K where the isentropic surface intersected a sloping tropopause, then the 2-PVU contour would be involved in lateral meanders. The corresponding version of Fig. 11 would then show filaments of tropospheric air being advected into the higher PV stratosphere and filaments of high PV stratospheric air similarly being advected into the troposphere. The alternating filaments of stratospheric and tropospheric air appear like interleaved fingers on an isentropic surface. Much of this exchange would be irreversible, given the rapid straining of features by the background wind field. There is observational evidence for stratosphere–troposphere exchange by waves of short vertical wavelength that produce interleaving laminae of tropospheric and stratospheric air at and above the tropospheric jet (Shapiro et al. 1980). Danielsen et al. (1991) and Kritz et al. (1991) show evidence of such filaments mixing air along isentropes in the stratosphere above the jet stream.

Aside from tracer transport, IGWs are important in that they constitute a subsynoptic signal having large amplitude in the vicinity of jet streaks and jet stream exit regions. Van Tuyl and Young (1982) emphasized this aspect of IGWs and the importance of not suppressing them during the process of assimilating new data into numerical forecast models. IGWs can also play a role in organizing convection and precipitation. Koch and Dorian (1988) observed IGWs from a jet streak exit region, organizing a sequence of thunderstorms as they propagated through a region of weak conditional convective stability.

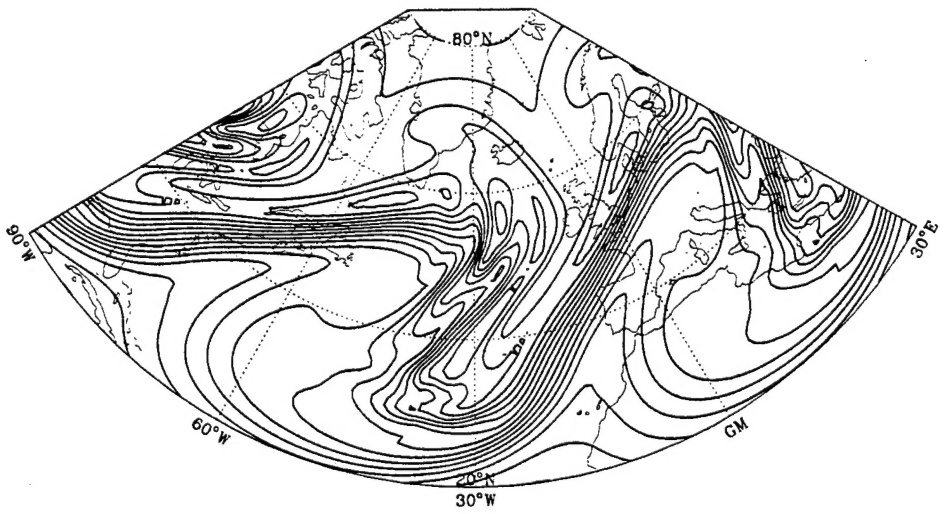


FIG. 11. Isentropic potential vorticity on the 430-K surface on day 11. Contour interval is 1 PVU ($\equiv 10^{-6}$ m^2 s^{-1} K kg^{-1}). This isentropic surface is near 90 mb at the location of the IGW packet.

/ams v3628 1203 Mp 19 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203

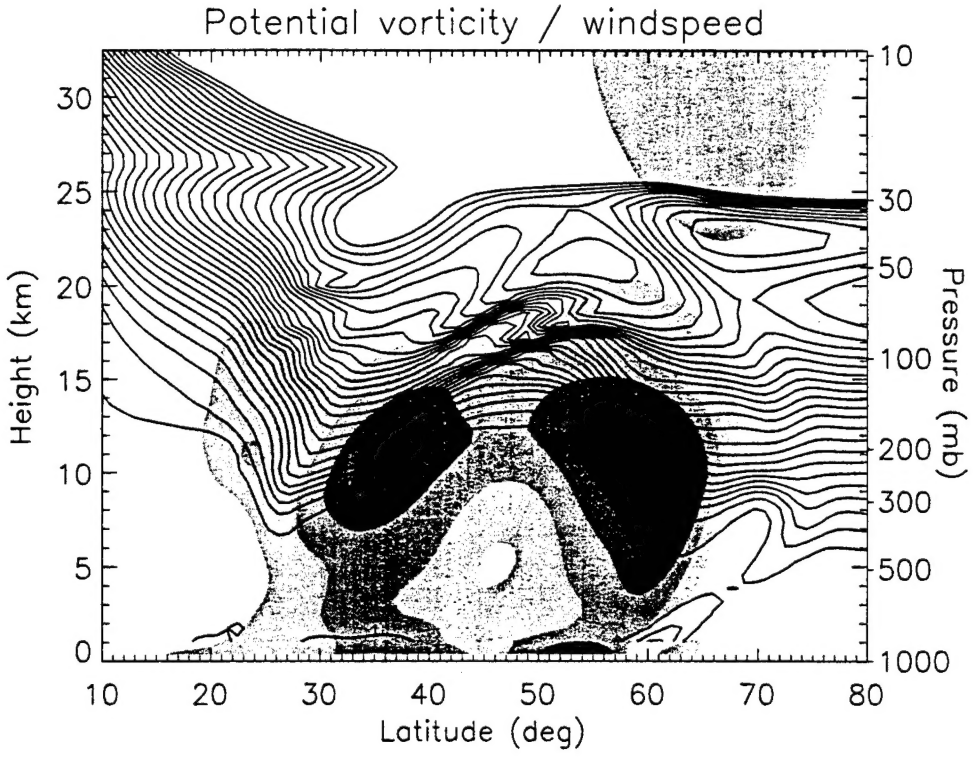


FIG. 12. Latitude–height cross section of PV at 30°W intersecting the lateral meanders of Fig. 11. Contour interval is 1 PVU.

Lastly, the upward momentum flux due to IGWs is considered. Given the widespread nature of IGWs and their preferred location in jetstreams, it might be thought that their upward flux of momentum could be important. However, the integrated momentum flux is proportional to energy times a factor $\delta_-^{1/2} = \sqrt{1-(f/\hat{\omega})^2}$, showing that smaller scale high-frequency gravity waves are more efficient in fluxing momentum vertically than IGWs (Fritts and van Zandt 1993). (The factor δ_+ in their formula is extraneous to the momentum flux and may be safely omitted.) It is of interest to note that the Eliassen–Palm flux approaches zero more rapidly as $\hat{\omega} \to f$, by an additional factor of δ_-. At the center of the wavepacket on day 11, $u'w' \approx -0.08\ \mathrm{m^2\ s^{-2}}$; a time- or area-average flux would be even smaller. Both are small numbers compared to the momentum flux near the tropopause estimated by Fritts and van Zandt.

IGWs could, nevertheless, have some impact on momentum transport by modulation of higher-frequency gravity waves (Broutman and Young 1986). Such waves might be absorbed at IGW-induced critical levels or, conversely, escape to higher altitudes. Partial reflection due to steep gradients created by IGW mixing is another possibility. Neither would be very important in the simulation where IGWs account for a small fraction of the total wind.

5. Conclusions

The excitation and propagation of inertia–gravity waves (IGWs) generated by an unstable baroclinic wave was examined with a high-resolution 3D nonlinear numerical model. IGWs arose spontaneously as the tropospheric jet stream was distorted by baroclinic instability and strong parcel accelerations took place, primarily in the jet stream exit region of the upper troposphere. Subsequent propagation of IGWs occurred in regions of strong windspeed, in the tropospheric and stratospheric jets, and in a cutoff low formed during the baroclinic lifecycle. IGWs on the flanks of these jets were rotated inward by differential advection and subsequently absorbed by the model's hyperdiffusion. This was the principal mechanism of IGW absorption in the model, the results being insensitive to vertical resolution. IGWs in reality are thought to break via Kelvin–Helmholtz (KH) instability, a process unresolved in the 3D model. Although the model's absorption of IGWs at the sidewalls of the jet is artificial, IGW propagation was for the most part confined to regions of $\hat{\omega} > f$. Only a few IGWs were able to penetrate the middle stratosphere, due to weak winds or an unfavorable alignment of wavevector with respect to the mean flow (except beneath the polar night jet).

IGWs are important both as a synoptic signal in the jet stream that may influence subsequent tropospheric

/ ams v3628 1203 Mp 20 Tuesday Jul 18 11:42 AM LP: J At Sci (August 15 95) 1203

developments and as a source of isentropic or cross-isentropic mixing in the lower stratosphere. Our results demonstrated for the first time numerically a significant isentropic displacement of potential vorticity isopleths due to IGWs above the tropopause. Since conditions for IGW propagation are favorable within a jet, a region of strong isentropic PV gradient, it is likely that inertia–gravity waves affect the permeability of the lower stratospheric vortex and may in some instances lead to stratosphere–troposphere exchange, particularly if IGW-induced parcel displacements are large and irreversible. Although IGW *breaking* may help ensure that the mixing is permanent, it is interesting to note from the simulations that straining by the large-scale flow, without IGW breaking, was sufficient to guarantee irreversible mixing of PV.

Observations of tracer laminae in the lower stratosphere sometimes suggest a concurrent IGW signal, while others do not; it is possible that "fossil" laminae were due to an earlier IGW event or, perhaps, the baroclinic instability itself. The large-scale shear distends tracer isopleths into many thin filaments (e.g., Plumb et al. 1994). It will be interesting to compare mixing rates with and without IGWs to assess their role in isentropic transport.

Considerably higher vertical (and horizontal) resolution is required to determine the mechanisms of IGW breaking and the role of IGWs in cross-isentropic transport. This issue is important in the lower stratosphere where other mechanisms of cross-isentropic transport are quite weak. The chemistry of this region is sensitive to the rate at which constituents are diffused vertically and exchanged between the troposphere and stratosphere.

IGW-induced mixing is expected to vary as a function of location and season. The climatology of IGWs and their relation to tropospheric sources are essentially unknown. If progress is to made, it will begin with an analysis of IGWs in relation to the synoptic-scale flow at locations where continuous radar or lidar observations of high-frequency motions are routinely available. Relationships established locally may then, as a first approximation, be extrapolated to global models.

Acknowledgments. Conversations with Kaoru Sato and Dick Reed are gratefully acknowledged. This research was supported by National Aeronautics and Space Administration, Contract NAS1-19563 and by Air Force Office of Scientific Research Contract F49620-92-C-0033.

REFERENCES

Barat, J., 1983: The fine structure of the stratospheric flow revealed by differential sounding. *J. Geophys. Res.*, **88,** 5219–5228.

Bénard, P., J.-L. Redelsperger, and J.-P. Lafore, 1992a: Nonhydrostatic simulation of frontogenesis in a moist atmosphere. Part I: General description and narrow rainbands. *J. Atmos. Sci.*, **49,** 2200–2217.

——, J.-P. Lafore, and J.-L. Redelsperger, 1992b: Nonhydrostatic simulation of frontogenesis in a moist atmosphere. Part II: Moist potential vorticity budget and wide rainbands. *J. Atmos. Sci.*, **49,** 2218–2235.

Blumen, W., 1972: Geostrophic adjustment. *Rev. Geophys. Space Phys.*, **10,** 485–528.

——, 1992: A semigeostrophic eady-wave frontal model incorporating momentum diffusion. Part III: Wave dispersion and dissipation. *J. Atmos. Sci.*, **49,** 1061–1074.

Broutman, D., and W. R. Young, 1986: On the interaction of small-scale oceanic internal waves with near-inertial waves. *J. Fluid Mech.*, **166,** 341–358.

Cammas, J.-P., and D. Ramond, 1989: Analysis and diagnosis of the composition of ageostrophic circulations in jet-front systems. *Mon. Wea. Rev.*, **117,** 2447–2462.

Chan, K. R., S. G. Scott, S. W. Bowen, S. E. Gaines, E. F. Danielsen, and L. Pfister, 1991: Horizontal wind fluctuations in the stratosphere during large-scale cyclogenesis. *J. Geophys. Res.*, **96,** 17 425–17 432.

Danielsen, E. F., R. S. Hipskind, W. L. Starr, J. F. Vedder, S. E. Gaines, D. Kley, and K. K. Kelly, 1991: Irreversible transport in the stratosphere by internal waves of short vertical wavelength. *J. Geophys. Res.*, **96,** 17 433–17 452.

Duffy, D. G., 1990: Geostrophic adjustment in a baroclinic atmosphere. *J. Atmos. Sci.*, **47,** 457–473.

Dunkerton, T. J., 1984: Inertia–gravity waves in the stratosphere. *J. Atmos. Sci.*, **41,** 3396–3404.

——, and R. E. Robins, 1992: Radiating and nonradiating modes of secondary instability in a gravity wave critical layer. *J. Atmos. Sci.*, **49,** 2546–2559.

Eckermann, S. D., and R. A. Vincent, 1993: VHF radar observations of gravity-wave production by cold fronts over southern Australia. *J. Atmos. Sci.*, **50,** 785–806.

Fritts, D. C., 1984: Gravity wave saturation in the middle atmosphere: A review of theory and observations. *Rev. Geophys. Space Phys.*, **22,** 275–308.

——, and P. K. Rastogi, 1985: Convective and dynamical instabilities due to gravity wave motions in the lower and middle atmosphere: Theory and observations. *Radio Sci.*, **20,** 1247–1277.

——, and G. D. Nastrom, 1992: Sources of mesoscale variability of gravity waves. Part II: Frontal, convective, and jetstream excitation. *J. Atmos. Sci.*, **49,** 111–127.

——, and T. E. van Zandt, 1993: Spectral estimates of gravity wave energy and momentum fluxes. Part I: Energy dissipation, acceleration, and constraints. *J. Atmos. Sci.*, **50,** 3685–3694.

Gall, R. L., R. T. Williams, and T. L. Clark, 1988: Gravity waves generated during frontogenesis. *J. Atmos. Sci.*, **45,** 2204–2219.

Garner, S. T., 1989: Fully Lagrangian numerical solutions of unbalanced frontogenesis and frontal collapse. *J. Atmos. Sci.*, **46,** 717–739.

Hoskins, B. J., and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.*, **101,** 637–655.

Keyser, D., and M. A. Shapiro, 1986: A review of the structure and dynamics of upper-level frontal zones. *Mon. Wea. Rev.*, **114,** 452–499.

——, B. D. Schmidt, and D. G. Duffy, 1989: A technique for representing three-dimensional vertical circulations in baroclinic disturbances. *Mon. Wea. Rev.*, **117,** 2463–2494.

Kitchen, E. H., and M. E. McIntyre, 1980: On whether inertio–gravity waves are absorbed or reflected when their intrinsic frequency is doppler-shifted towards *f*. *J. Meteor. Soc. Japan,* **58,** 118–125.

Koch, S. E., and P. B. Dorian, 1988: A mesoscale gravity wave event observed during CCOPE. Part III: Wave environment and probable source mechanisms. *Mon. Wea. Rev.*, **116,** 2570–2592.

Kritz, M. A., S. W. Rosner, E. F. Danielsen, and H. B. Selkirk, 1991: Air mass origins and troposphere-to-stratosphere exchange associated with mid-latitude cyclogenesis and tropopause folding inferred from 7Be measurements. *J. Geophys. Res.*, **96,** 17 405–17 414.

Ledwell, J. R., A. J. Watson, and C. S. Law, 1993: Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature,* **364,** 701–703.

Lelong, M. P., and J. C. McWilliams, 1993: Cyclostrophic adjustment. *Phys. Fluids,* submitted.

Ley, B., and W. R. Peltier, 1978: Wave generation and frontal collapse. *J. Atmos. Sci.,* **35,** 3–17.

Lindzen, R. S., and M. Fox-Rabinovitz, 1989: Consistent vertical and horizontal resolution. *Mon. Wea. Rev.,* **117,** 2575–2583.

Luo, Z., and D. C. Fritts, 1993: Gravity-wave excitation by geostrophic adjustment of the jet stream. Part II: Three-dimensional forcing. *J. Atmos. Sci.,* **50,** 104–115.

Norton, W. A., 1994: Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours. *J. Atmos. Sci.,* **51,** 654–673.

Pierce, R. B., and T. D. A. Fairlie, 1993: Chaotic advection in the stratosphere: Implications for the dispersal of chemically perturbed air from the polar vortex. *J. Geophys. Res.,* **98,** 18 589–18 595.

Plumb, R. A., and Coauthors, 1994: Intrusions into the lower stratospheric Arctic vortex during the winter of 1991–1992. *J. Geophys. Res.,* **99,** 1089–1105.

Prichard, I. T., and L. Thomas, 1993: Radar observations of gravity-wave momentum fluxes in the troposphere and lower stratosphere. *Ann. Geophys.,* **11,** 1075–1083.

Sato, K., 1989: An inertial gravity wave associated with a synoptic-scale pressure trough observed by the MU radar. *J. Meteor. Soc. Japan,* **67,** 325–334.

——, 1993: Small-scale wind disturbances observed by the MU radar during the passage of typhoon Kelly. *J. Atmos. Sci.,* **50,** 518–537.

——, 1994: A statistical study of the structure, saturation and sources of inertio–gravity waves in the lower stratosphere observed with the MU radar. *J. Atmos. Terr. Phys.,* **56,** 755–774.

Sato, T., and R. F. Woodman, 1982: Fine altitude resolution radar observations of upper-tropospheric and lower-stratospheric winds and waves. *J. Atmos. Sci.,* **39,** 2539–2545.

Shapiro, M. A., 1980: Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the stratosphere and troposphere. *J. Atmos. Sci.,* **37,** 994–1004.

——, and D. Keyser, 1990: Fronts, jetstreams, and the tropopause. *Extratropical Cyclones: The Erik Palmén Memorial Volume,* C. W. Newton and E. O. Holopainen, Eds., Amer. Meteor. Soc., 262 pp.

——, E. R. Reiter, R. D. Cadle, and W. A. Sedlacek, 1980: Vertical mass and trace constituent transports in the vicinity of jet streams. *Arch. Meteor. Geophys. Bioklimatol., Ser. B,* **28,** 193–206.

Shepherd, T. G., 1990: A general method for finding extremal states of Hamiltonian dynamical systems, with applications to perfect fluids. *J. Fluid Mech.,* **213,** 573–587.

Simmons, A. J., and B. J. Hoskins, 1978: The lifecycles of some nonlinear baroclinic waves. *J. Atmos. Sci.,* **35,** 414–432.

——, and ——, 1980: Barotropic influences on the growth and decay of nonlinear baroclinic waves. *J. Atmos. Sci.,* **37,** 1679–1684.

Snyder, C., W. C. Skamarock, and R. Rotunno, 1993: Frontal dynamics near and following frontal collapse. *J. Atmos. Sci.,* **50,** 3194–3212.

Thomas, L., I. T. Prichard, and I. Astin, 1992: Radar observations of an inertia–gravity wave in the troposphere and lower stratosphere. *Ann. Geophys.,* **10,** 690–697.

Thorncroft, C. D., and B. J. Hoskins, 1990: Frontal cyclogenesis. *J. Atmos. Sci.,* **47,** 2317–2336.

Uccellini, L. W., and S. E. Koch, 1987: The synoptic setting and possible energy sources for mesoscale wave disturbances. *Mon. Wea. Rev.,* **115,** 721–729.

Ushimaru, S., and H. Tanaka, 1990: Characteristics of internal inertial gravity waves and inertial waves in the lower stratosphere observed by the MU radar. *J. Meteor. Soc. Japan,* **68,** 1–18.

Vallis, G. K., 1992: Mechanisms and parameterizations of geostrophic adjustment and a variational approach to balanced flow. *J. Atmos. Sci.,* **49,** 1144–1160.

——, G. F. Carnevale, and W. R. Young, 1989: Extremal energy properties and construction of stable solutions of the Euler equations. *J. Fluid Mech.,* **207,** 133–152.

Van Tuyl, A. H., and J. A. Young, 1982: Numerical simulation of nonlinear jet streak adjustment. *Mon. Wea. Rev.,* **110,** 2038–2054.

Warn, T., and R. Menard, 1986: Nonlinear balance and gravity–inertial wave saturation in a simple atmospheric model. *Tellus,* **38A,** 285–294.

Waugh, D. W., and R. A. Plumb, 1994: Contour advection with surgery: A technique for investigating finescale structure in tracer transport. *J. Atmos. Sci.,* **51,** 530–540.

——, and Coauthors, 1994: Transport out of the lower stratospheric Arctic vortex by Rossby wave breaking. *J. Geophys. Res.,* **99,** 1071–1088.

Yamanaka, M. D., and H. Tanaka, 1984: Multiple 'gust layers' observed in the middle stratosphere. *Dynamics of the Middle Atmosphere,* J. R. Holton and T. Matsuno, Eds., Terra Scientific, 117–140.

Young, R. E., and G. L. Villere, 1985: Nonlinear forcing of planetary scale waves by amplifying unstable baroclinic eddies generated in the troposphere. *J. Atmos. Sci.,* **42,** 1991–2006.